Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **418,** 2074–2105 (2011) doi:10.1111/j.1365-2966.2011.19626.x

A robust sample of galaxies at redshifts $6.0 < z < 8.7$: stellar populations, star formation rates and stellar masses

R. J. McLure,[1⋆] J. S. Dunlop,[1] L. de Ravel,[1] M. Cirasuolo,[1,2] R. S. Ellis,[3] M. Schenker,[3] B. E. Robertson,[3] A. M. Koekemoer,[4] D. P. Stark[5] and R. A. A. Bowler[1]

[1]*SUPA,† Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ*
[2]*Astronomy Technology Centre, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ*
[3]*Department of Astronomy, California Institute of Technology, Pasadena, CA 91125, USA*
[4]*Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA*
[5]*Kavli Institute of Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA*

Accepted 2011 August 11. Received 2011 July 4; in original form 2011 February 28

ABSTRACT

We present the results of a photometric redshift analysis designed to identify $z \geq 6$ galaxies from the near-infrared *Hubble Space Telescope* imaging in three deep fields [*Hubble Ultra Deep Field* (HUDF), HUDF09-2 and Early Release Science] covering a total area of 45 square arcmin. By adopting a rigorous set of criteria for rejecting low-redshift interlopers, and by employing a deconfusion technique to allow the available ultradeep IRAC imaging to be included in the candidate-selection process, we have derived a robust sample of 70 Lyman break galaxies (LBGs) spanning the redshift range $6.0 < z < 8.7$. Based on our final sample, we investigate the distribution of ultraviolet (UV) spectral slopes ($f_\lambda \propto \lambda^\beta$), finding a variance-weighted mean value of $\langle\beta\rangle = -2.05 \pm 0.09$ which, contrary to some previous results, is not significantly bluer than displayed by lower redshift starburst galaxies. We confirm the correlation between UV luminosity and stellar mass reported elsewhere, but based on fitting galaxy templates featuring a range of star formation histories (SFHs), metallicities and reddening, we find that, at $z \geq 6$, the range in mass-to-light ratio ($M_\star/L_{\rm UV}$) at a given UV luminosity could span a factor of $\simeq$50. Focusing on a subsample of 21 candidates with IRAC detections at 3.6 μm, we find that $L^\star$ LBGs at $z \simeq 6.5$ have a median stellar mass of $M_\star = (2.1 \pm 1.1) \times 10^9\,{\rm M}_\odot$ (Chabrier initial mass function) and a median specific star formation rate (sSFR) of $1.9 \pm 0.8\,{\rm Gyr}^{-1}$. Using the same subsample, we have investigated the influence of nebular continuum and line emission, finding that for the majority of candidates (16 out of 21), the best-fitting stellar masses are reduced by less than a factor of 2.5. However, galaxy template fits exploring a plausible range of SFHs and metallicities provide no compelling evidence of a clear connection between SFR and stellar mass at these redshifts. Finally, a detailed comparison of our final sample with the results of previous studies suggests that, at faint magnitudes, several high-redshift galaxy samples in the literature are significantly contaminated by low-redshift interlopers.

Key words: galaxies: evolution – galaxies: formation – galaxies: high-redshift.

1 INTRODUCTION

The goal of identifying and studying the nature of ultrahigh-redshift galaxies remains one of the most important challenges in observational cosmology, and holds the key to furthering our understanding of the earliest stages of galaxy evolution and unveiling the nature of the sources responsible for cosmic reionization.

Observational constraints provided by the Gunn–Peterson trough in the spectra of high-redshift quasars (e.g. Fan et al. 2006) suggest that reionization was coming to an end at $z \simeq 6$ (Becker, Rauch & Sargent 2007). Moreover, optical polarization measurements from the *WMAP* experiment indicate that reionization began at $z \simeq 11$ if it is assumed to be a single, rapid event (Dunkley et al. 2009). Consequently, it is now apparent that to improve our understanding of cosmic reionization and to unveil the earliest epoch of galaxy formation, it is necessary to extend studies of high-redshift galaxies into the $7 < z < 10$ régime (e.g. Robertson et al. 2010).

⋆E-mail: rjm@roe.ac.uk
†Scottish Universities Physics Alliance.

Given our existing knowledge of the evolution of the galaxy luminosity function in the redshift interval $5.0 < z < 6.5$ (e.g. Bouwens et al. 2007; McLure et al. 2009), it is clear that achieving this aim requires ultradeep near-infrared (near-IR) imaging, reaching detection limits of $26 < m_{\rm AB} < 30$. At the bright end of this range, wide-field, ground-based imaging has a unique contribution to make, and has recently allowed the luminosity function, the clustering properties and the stellar populations of luminous Lyman break galaxies (LBGs: McLure et al. 2009; Ouchi et al. 2009; Grazian et al. 2011) and Lyman alpha emitters (LAEs: Ono et al. 2010; Ota et al. 2010; Ouchi et al. 2010; Nakamura et al. 2011) to be studied in detail. Indeed, the importance of ground-based imaging and spectroscopy has recently been highlighted by the spectroscopic confirmation of two LBGs at $z = 7.01$ and 7.11 by Vanzella et al. (2010). However, in advance of 30-m-class ground-based telescopes, it is clear that routinely identifying and studying sub-$L^\star$ galaxies at $z \geq 7$ is only possible using space-based imaging.

Consequently, the unparalleled near-IR sensitivity provided by the new WFC3 camera, installed on the *Hubble Space Telescope* (*HST*) in late 2009, has proven to be a crucial breakthrough in high-redshift galaxy studies. Indeed, despite only covering an area of $\simeq$ 4.5 square arcmin, the unprecedented depth ($m_{\rm AB} \simeq 29$, 5σ) of the first tranche of WFC3/IR imaging of the *Hubble Ultra Deep Field* (HUDF; GO-11563) led to a raft of early science papers investigating the number densities, luminosity functions, stellar masses and stellar populations of $6.5 < z < 8.5$ galaxies (e.g. Bouwens et al. 2010a,b; Bunker et al. 2010; Finkelstein et al. 2010; Labbé et al. 2010; McLure et al. 2010; Oesch et al. 2010; Yan et al. 2010). Interestingly, Lehnert et al. (2010) have recently claimed the tentative detection of Lyman α emission at $z = 8.56$ in a WFC3/IR candidate in the HUDF, originally identified by Bouwens et al. (2010a) and McLure et al. (2010). Although the large rest-frame equivalent width (EW $\simeq$ 210 Å) of the Lyman α emission line suggests that, if confirmed, this object must be a decidedly atypical example of a $z \geq 6$ LBG (Stark et al. 2010), the location of the claimed Lyman α emission line is in good agreement with the original photometric redshifts derived by McLure et al. (2010) and Finkelstein et al. (2010): $z_{\rm phot} = 8.45 \pm 0.50$ and 8.61 ± 0.35, respectively.

In addition to the WFC3/IR imaging of the HUDF, another key data set has been the WFC3/IR imaging taken as part of the Early Release Science extragalactic programme (ERS; GO-11359) which, although substantially shallower than the HUDF WFC3/IR imaging ($m_{\rm AB} \simeq 27.5$, 5σ), covers an area approximately 10 times larger. By combining the HUDF and ERS data sets to obtain a greater dynamic range in ultraviolet (UV) luminosity, Bouwens et al. (2010b) and Labbé et al. (2010) investigated the relationship between the UV spectral slope ($f_\lambda \propto \lambda^\beta$) and UV luminosity. Both studies find a correlation, with β changing from $\beta \simeq -2$ (a typical value for lower redshift starburst galaxies) at $M_{1500} \simeq -20.5$ to extremely blue values of $\beta \simeq -3$ at $M_{1500} \simeq -18.5$. As discussed by Bouwens et al. (2010b), although dust-free, low-metallicity models can produce slopes of $\beta \simeq -3$, they can only do so under the assumption that the ionizing photon escape fraction is high ($f_{\rm esc} > 0.3$) and that, correspondingly, the contribution from nebular continuum emission is low.

Based on stacking the ACS+WFC3/IR+IRAC photometry of $z \simeq 7$ LBG candidates in the HUDF and ERS data sets, Labbé et al. (2010) find the same correlation between M_{1500} and spectral slope β as Bouwens et al. (2010b). However, Labbé et al. (2010) conclude that it is not possible to reproduce both the blue spectral slopes and significant $\lambda_{\rm rest} \simeq 4000$ Å spectral breaks displayed by the faintest $z \simeq 7$ LBG candidates, without recourse to episodic star formation histories (SFHs) and/or a significant contribution from nebular line emission. Indeed, Ono et al. (2010) also conclude that nebular line emission may be necessary to reproduce the observed $J - m_{3.6}$ colour in a stack of $z \simeq 6$ LAE photometry.

Despite the uncertainties, one observational result that has received significant attention recently is the apparent relationship between star formation rate (SFR) and stellar mass. Both Labbé et al. (2010) and González et al. (2010) find an approximately linear correlation between stellar mass and SFR at $z \simeq 7$, consistent with the results derived by Stark et al. (2009) for LBGs in the redshift range $4 < z < 6$. As a result, Labbé et al. (2010) and González et al. (2010) conclude that the specific SFR (sSFR) of $z \simeq 7$ LBGs is remarkably constant (sSFR $\simeq 2$ Gyr^{-1}), and consistent with the value of sSFR $\simeq 2.5$–4.5 Gyr^{-1} observed in star-forming galaxies at $z \simeq 2$–3 (e.g. Daddi et al. 2007; Magdis et al. 2010). As previously discussed by Stark et al. (2009), a natural explanation of this observation would be to invoke a SFR which exponentially increases with time although, as shown by Finlator, Oppenheimer & Davé (2011), SFHs of this type may have difficulty in reproducing some of the most extreme Balmer breaks reported in the literature at $z \geq 6$.

The majority of previous studies which have investigated the high-redshift galaxy population using WFC3/IR imaging have relied on traditional colour-cut, or 'drop-out', selection techniques. In contrast, the principal motivation for this paper is to investigate what can be learned about the $z \geq 6$ galaxy population by fully exploiting the excellent multiwavelength (ACS+WFC3/IR+IRAC) data which are now available over an area of $\simeq$45 square arcmin. Rather than applying standard 'drop-out' criteria, in this work, we continue to pursue the strategy we have previously adopted (McLure et al. 2006, 2009, 2010) and employ a template-fitting, photometric redshift analysis to select our final high-redshift galaxy sample. A key new element in this strategy is our development of a deconfusion algorithm capable of providing the robust IRAC photometry necessary for improved photometric redshift and stellar-mass estimates.

In principle, this technique should have several advantages over the standard LBG 'drop-out' selection. First, by employing all of the available multiwavelength data, including the IRAC photometry, it is possible to make optimal use of the available information. Secondly, by avoiding any colour pre-selection, this approach should be less biased towards simply selecting the very bluest galaxies at high redshift. This second point is potentially crucial in the context of investigating the claims of ultrablue UV spectral slopes for LBG candidates at $z \geq 6$ (see Section 4). Finally, a spectral energy distribution (SED) fitting analysis also provides an estimate of the photometric redshift probability density function, $P(z)$, and therefore allows the prevalence, and significance, of competing photometric redshift solutions at low redshift to be transparently investigated.

The primary motivation of this paper is therefore to construct the most robust sample possible using the available data and techniques reviewed above, in order to critically address some of the claims about the properties of the $z > 6$ population newly found with the *HST*. The structure of this paper is as follows. In Section 2, we describe the available data in each of the three fields, including a brief description of our IRAC deconfusion algorithm. In Section 3, we describe our initial candidate-selection, photometric redshift analysis and the construction of the final catalogue of robust candidates. In Sections 4 and 5, we investigate the UV spectral slopes, stellar masses and SFRs of the final robust sample. In Section 6, we perform a detailed comparison of our final robust sample with samples previously derived in the literature, exploring the reasons behind any apparent discrepancies. In Section 7, we provide a summary of our

Table 1. The results of the image depth analysis described in Section 3.2. Columns 1–4 list the survey field names, central coordinates and areas (in square arcmin). The areas listed in column 4 correspond to the areas of WFC3/IR imaging which were actually useful for the analysis performed in this paper (i.e. excluding array edges and areas which are not covered by ACS optical imaging). The remaining columns list the median 5σ depths in each available ACS+WFC3/IR+IRAC filter. The depths are all referenced to a 0.6-arcsec-diameter aperture, and are not aperture-corrected. Due to the fact that the IRAC depths are determined via a deconfusion process (see Section 2.3.1), the local IRAC depths can differ significantly from the median values listed here.

Field	RA (J2000)	Dec. (J2000)	Area	B_{435}	V_{606}	i_{775}	z_{850}	Y_{098}	Y_{105}	J_{125}	H_{160}	3.6 μm	4.5 μm
HUDF	03:32:38.5	−27:46:57.0	4.5	29.04	29.52	29.19	28.54	–	28.59	28.67	28.73	26.3	25.9
HUDF09-2	03:32:23.4	−27:42:52.0	4.5	–	28.49	28.22	28.06	–	28.24	28.60	28.49	26.2	–
ERS	03:33:05.5	−27:51:21.6	36.5	27.68	27.87	27.29	27.06	27.26	–	27.66	27.40	26.0	25.6

main conclusions. In the appendices, we provide a full description of our IRAC deconfusion procedure, full photometry and grey-scale postage-stamp images for each high-redshift candidate and individual plots illustrating the results of our SED fitting. All magnitudes are quoted in the AB system (Oke & Gunn 1983) and all calculations assume $\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

2 DATA

The analysis in this paper relies on the publicly available optical, near-IR and mid-IR imaging data covering the HUDF, HUDF09-2 and ERS field. In this section, we briefly describe the relevant details of the various imaging data sets, and the depth analysis which was performed in order to attribute accurate error estimates to the candidate photometry. The basic properties of the three fields are listed in Table 1.

2.1 WFC3/IR imaging

The WFC3/IR imaging of both the HUDF and the HUDF09-2 was taken as part of the public treasury programme GO-11563 (PI: G. Illingworth)[1] and consists of single pointings ($\simeq$4.5 square arcmin) of the WFC3/IR in the *F*105*W*, *F*125*W* and *F*160*W* filters (hereinafter referred to as Y_{105}, J_{125} and H_{160}, respectively). The WFC3/IR data set in the ERS field was taken as part of the public programme GO-11359 (PI: O'Connell) and consists of a mosaic of 10 pointings of the WFC3/IR in the *F*098*M* (Y_{098}), J_{125} and H_{160} filters (Windhorst et al. 2011).[2] The WFC3/IR data were calibrated using CALWF3 and subsequently combined using MULTIDRIZZLE (Koekemoer et al. 2002) as summarized in (McLure et al. 2010); full details are presented in Koekemoer et al. (2011). The final mosaics have point spread functions (PSFs) with full width at half-maximum (FWHM) in the range 0.15–0.18 arcsec depending on the filter, and were drizzled on to a final grid of 0.06 arcsec pixel^{-1}. In the case of the HUDF and ERS mosaics, the final astrometry was matched to that of the publicly available reductions of the optical ACS imaging of the UDF (Beckwith et al. 2006) and GOODS-S (GOODSv2.0; Giavalisco et al. 2004), respectively. The astrometry for the final mosaics of the HUDF09-2 was matched to the *J*-band imaging of the GOODS-S taken as part of the MUSYC survey (Cardamone et al. 2010), with a typical rms accuracy of $\simeq$0.1 arcsec. The WFC3/IR imaging of the ERS analysed in this paper consists of the data comprising the completed programme. However, for the HUDF09-2 and HUDF, we make use of the epoch 1 observations, which consist of the data publicly available as of 2010 February and 2010 August, respectively.

2.2 ACS imaging

For the HUDF and ERS field, the ACS data used in this study consists of the publicly available reductions of the *F*435*W*, *F*606*W*, *F*775*W* and *F*850*LP* (hereinafter B_{435}, V_{606}, i_{775} and z_{850}, respectively) imaging of the HUDF (Beckwith et al. 2006) and GOODS-S (GOODSv2.0; Giavalisco et al. 2004). The ACS imaging covering the HUDF09-2 is our own reduction (based on CALACS and MULTIDRIZZLE) of the V_{606}, i_{775} and z_{850} imaging obtained as part of the UDF05 programme (Oesch et al. 2007). All of the optical ACS imaging was re-sampled to a 0.06 arcsec pixel^{-1} grid to match the WFC3/IR data, and the astrometry of the ACS data covering the HUDF09-2 was also registered to match the MUSYC imaging of the GOODS-S.

2.3 IRAC imaging

For the HUDF and ERS field, we make use of the publicly available reductions (v0.30) of the 3.6- and 4.5-μm IRAC imaging obtained as part of the GOODS (proposal ID 194, Dickinson et al., in preparation). The IRAC data covering the ERS consist of approximately 23 hours of on-source integration in both the 3.6- and 4.5-μm bands. For the HUDF, we performed an inverse-variance-weighted stack of the overlapping region of the epoch 1 and epoch 2 imaging, producing final mosaics consisting of approximately 46 hours of on-source integration at 3.6 and 4.5 μm. For the HUDF09-2, we re-registered and stacked the MOPEX reductions of the 3.6-μm imaging obtained via proposal ID 30866 (PI: R. Bouwens) producing a final 3.6-μm mosaic with an on-source integration time of approximately 33 hours. The consistency of the IRAC photometry was checked via reference to the SIMPLE (Damen et al. 2011) imaging at 3.6 and 4.5 μm of the *Extended Chandra Deep Field*-South, which overlaps all three fields. The astrometry of the IRAC imaging in all three fields was registered to that of the corresponding H_{160} WFC3/IR imaging.

2.3.1 *IRAC deconfusion*

A key new feature of the analysis undertaken in this paper is the inclusion of the IRAC photometry in the candidate-high-redshift-galaxy-selection procedure. Due to its depth, and comparatively broad PSF (FWHM $\simeq$ 1.5 arcsec), the 3.6 + 4.5 μm IRAC imaging covering the three fields of interest is heavily confused, making the process of obtaining aperture photometry matched to the optical/near-IR *HST* imaging non-trivial. In order to achieve this aim, it is therefore necessary to pursue some form of deconfusion

[1] We do not consider the third WFC3/IR pointing obtained as part of GO-11563, HUDF09-1, because deep IRAC imaging of this field is not currently available.

[2] All of the WFC3/IR data utilized in this paper conform to the nominal flight zero-points, that is, $Y_{098M} = 25.68$, $Y_{105W} = 26.27$, $J_{125W} = 26.25$ and $H_{160W} = 25.96$.

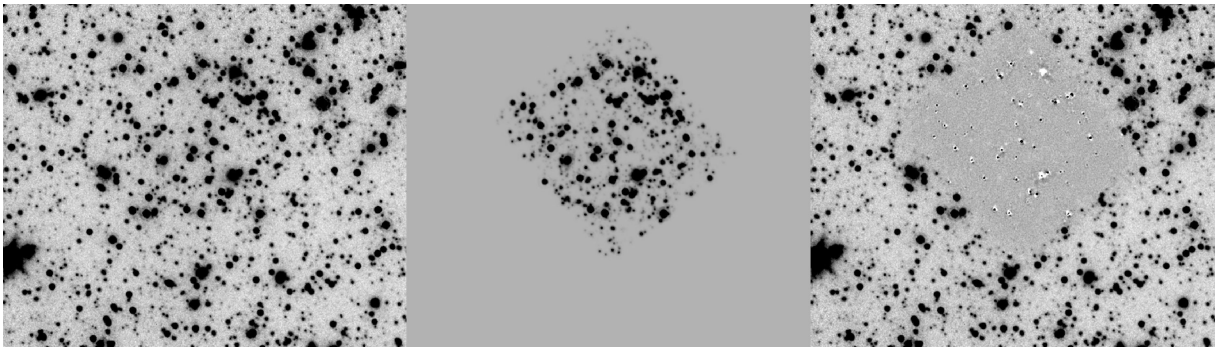

Figure 1. Illustration of the IRAC deconfusion algorithm. The left-hand panel shows the inverse-variance-weighted stack of the epoch 1+epoch 2 4.5-μm imaging covering the HUDF. The middle panel shows the best-fitting model of the IRAC data, based on using the H_{160} WFC3/IR imaging to provide model templates, and a matrix inversion procedure to determine the best-fitting template amplitudes (see Appendix A for full details). The right-hand panel shows the model subtracted image (note that the WFC3/IR imaging does not cover the full area of the HUDF).

process, which allows the IRAC imaging to be utilized beyond the natural confusion limit. Although there are several techniques which can be used to deconfuse IRAC imaging (see Appendix A), we have developed our own software which uses the WFC3/IR imaging data to provide normalized templates for each object in the field and then, via a transfer function, produces synthetic IRAC images on the native 0.6 arcsec pixel^{-1} plate scale. Through a matrix inversion procedure, the amplitude (or total flux) of each template can be simultaneously fitted to produce the optimal reproduction of the observed IRAC image (see Fig. 1). As a result of this procedure, it is effectively possible to extract accurate aperture photometry from the 3.6- and 4.5-μm IRAC imaging at the spatial resolution of the WFC3/IR imaging. An additional advantage of this approach is that it naturally provides robust uncertainties on the delivered flux measurements, which depend both on the signal-to-noise ratio of the IRAC detection *and* on the local level of confusion in the IRAC image.

3 CANDIDATE SELECTION

The process of candidate selection can be broken down into three separate stages: object detection and photometry, photometric redshift analysis and sample cleaning. Each stage in the process is described below.

3.1 Object detection and photometry

The initial catalogue construction process was identical in each of the three fields, and relied on SExtractor v2.5.0 (Bertin & Arnouts 1996). Preliminary catalogues were constructed in which object detection was performed in the Y_{105W}/Y_{098M}, J_{125W} and H_{160W} bands, using an aggressive set of SExtractor parameters, with matched photometry extracted from the corresponding ACS imaging by running SExtractor in dual-image mode. The separate catalogues were then concatenated to produce a master catalogue of unique objects in each of the three fields.

In order to avoid biases which can be introduced by adopting small photometric apertures, all of the analysis in this paper is based on 0.6-arcsec-diameter aperture photometry. For the purposes of the photometric redshift analysis, the fluxes from the 0.6-arcsec-diameter aperture photometry are not corrected to total, but the WFC3/IR and IRAC fluxes are corrected by small amounts (2–10 per cent) to account for aperture losses *relative* to the ACS imaging.

3.2 Depth analysis

A crucial part of the analysis necessary to identify robust high-redshift candidates is the derivation of accurate photometric uncertainties in each band. This was achieved by first producing a so-called χ^2 image (Szalay, Connolly & Szokoly 1999) of the registered optical+near-IR images of each field to identify which pixels are genuine 'blank sky'. Secondly, a grid of 0.6-arcsec-diameter apertures was placed in the blank-sky regions on each image. Thirdly, in order to determine the local image depth for each candidate, in each filter, the rms aperture-to-aperture variation was determined, by examining the distribution of the nearest 50 blank apertures. In this fashion, we are able to determine a local depth measurement for each individual candidate. For information, the median 5σ depths for each field are listed in Table 1.

3.3 Photometric redshift analysis

To perform the SED analysis necessary for this study, we have developed a new, bespoke, template-fitting code. The primary motivation for developing this new code was to provide the freedom to explore the relevant multidimensional parameter space in detail, investigating the impact of different SED templates, initial mass functions (IMFs), dust attenuation prescriptions and intergalactic medium (IGM) absorption recipes. Moreover, by employing our own software, it is possible to have full control over which derived quantities are provided as output, and the exact details of how the template fitting is performed. For example, our new code performs the SED fitting based on flux densities (f_ν), rather than magnitudes, which has the advantage of allowing the flux errors to be dealt with in a rigorous manner. Moreover, if necessary, the new code offers the possibility of fitting the input photometry with multicomponent stellar populations, each with separate metallicities and/or dust attenuation prescriptions.

For the purposes of this study, we employed the Bruzual & Charlot (2003) and Charlot & Bruzual (private communication) stellar evolution models (hereinafter BC03 and CB07, respectively), considering models with metallicities ranging from solar ($Z_\odot$) to 1/50th solar ($0.02\,Z_\odot$). Models with instantaneous bursts of star formation, constant star formation and SFRs exponentially declining with characteristic time-scales in the range $50\,\mathrm{Myr} < \tau < 10\,\mathrm{Gyr}$ were all considered. The ages of the stellar population models were allowed to range from 10 Myr to 13.7 Gyr, but were required to be less than the age of the Universe at each redshift. Dust reddening

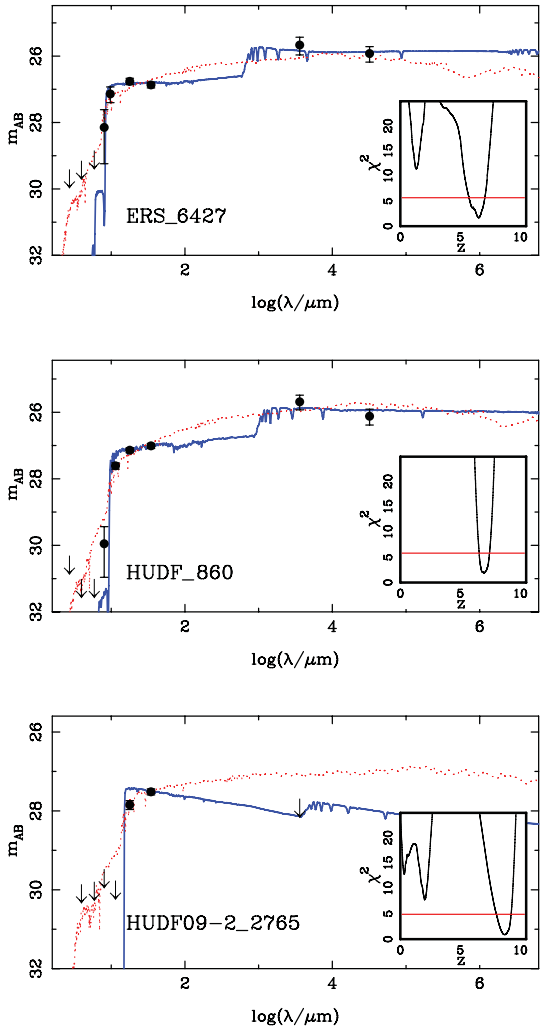

Figure 2. Example plots showing the results of our SED fitting for three objects from our final robust sample at redshifts $z_{\rm phot} = 6.5$, 7.0 and 8.7 (top to bottom). In each plot, the thick (blue) line shows the best-fitting high-redshift galaxy template, and the thin dotted (red) line shows the best-fitting alternative solution at low redshift. In each panel, the inset shows the value of χ^2 as a function of redshift (marginalized over all other free parameters). The horizontal line in each inset panel highlights the location of $(\chi^2_{\rm min} + 4)$ which corresponds to our requirement that a candidate is considered 'robust' only if any alternative low-redshift solutions can be ruled out at the ≥95 per cent confidence level. In each example, the combination of image depth and wavelength coverage allows us to rule out any alternative low-redshift solutions with high confidence. The bottom panel (HUDF09-2_2765) provides a good example of a situation where even a robust upper limit to the IRAC 3.6-μm flux provides a powerful redshift discriminant.

was described by the Calzetti et al. (2000) attenuation law, and allowed to vary within the range $0.0 < A_V < 2.5$ mag. IGM absorption shortwards of Lyman α was described by the Madau (1995) prescription, and a Chabrier (2003) IMF was assumed in all cases.[3] In Fig. 2, we show example SED fits for three objects (one from each field) covering the redshift range $6.5 < z_{\rm phot} < 8.7$.

[3] Derived quantities such as stellar masses and SFRs can be converted to a Salpeter (1955) IMF by multiplying by a factor of 1.8.

3.4 Sample cleaning

Based on the results of the photometric redshift fitting, all objects which displayed a statistically acceptable solution at $z_{\rm phot} \geq 4.5$ were retained, while those with no acceptable solution at high redshift were excluded. In each of the three fields, this initial screening process removed more than 90 per cent of the original input catalogues. Following the initial photometric redshift fitting, the remaining samples of potential high-redshift candidates were manually screened to remove artefacts (e.g. diffraction spikes), edge effects and spurious candidates such as high-surface-brightness features within the extended envelopes of luminous low-redshift galaxies.

3.4.1 Final candidate sample

From a practical perspective, the primary goal of this study is to produce a robust sample of high-redshift galaxy candidates at $z \geq 6$. In order to achieve this aim, three criteria were applied to the remaining potential high-redshift candidates:

(i) statistically acceptable redshift solution at $z_{\rm phot} \geq 6.0$;
(ii) secondary redshift solution excluded at ≥95 per cent confidence;
(iii) integrated probability $\int_{z=6}^{z=10} P(z')\delta z' \geq 0.5$.

The first criterion simply restricts our final sample to those objects for which the best-fitting SED template lies at $z_{\rm phot} \geq 6$. The second criterion rejects those objects for which the competing low-redshift solution cannot be excluded at high confidence. Specifically, this criterion is enforced by insisting that the $\Delta\chi^2$ between the primary and secondary photometric redshift solutions (following marginalization over all other relevant parameters) is ≥4. The final criterion is designed to exclude a small number of candidates with relatively flat $P(z)$ distributions for which, despite having a primary photometric redshift solution at $z_{\rm phot} \geq 6$, the majority of their integrated probability density distribution function lies at $z_{\rm phot} \leq 6$. We note that this final criterion is very similar to that employed by Finkelstein et al. (2010) in their HUDF analysis.

Our final robust sample of $z \geq 6$ galaxy candidates consists of $N = 70$ objects, spanning the redshift range $6.0 < z < 8.7$ and covering more than a factor of 10 in intrinsic UV luminosity, from $-18.2 < M_{1500} < -21.2$.[4] It is perhaps worth pointing out that if we had only insisted on a statistically acceptable primary photometric redshift solution at $z_{\rm phot} \geq 6.0$, the final sample would have contained $N = 130$ candidates. It should be stressed that it is likely that a significant fraction of the excluded objects are indeed $z \geq 6$ galaxies (see Fig. 3); it is simply that with the data in hand, it is not possible to consider them as robust candidates.

Due to the fact that LBGs at $z \geq 6$ are necessarily young galaxies, the differences between the SED fits provided by the BC03 and CB07 models are negligible. Consequently, to ease the comparison with previous studies, throughout the rest of this paper we adopt the results of the SED-fitting analysis based on the BC03 models. The final robust samples in each of the three fields, along with the best-fitting photometric redshift solutions and various other derived parameters, are presented in Tables 2–4. Note that, although we make no further use of this information throughout the rest of this paper, in Tables 2–4, we also list the best-fitting photometric redshift

[4] The final sample consists of 73 objects if three additional objects are included which satisfy our selection criteria only when Lyman α emission is included in the SED templates.

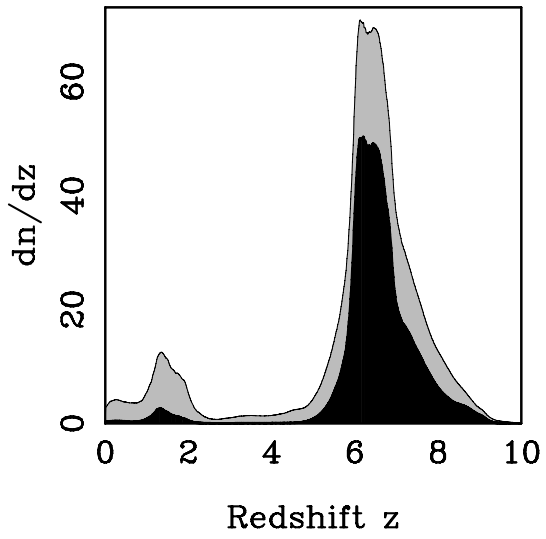

Figure 3. In black, we show the redshift distribution of our final robust sample, which has been calculated by summing the estimated redshift probability density function of each candidate. This is our best estimate of the redshift distribution of the $N = 70$ objects which satisfy all three of the criteria listed in Section 3.4.1. In grey, we show the redshift distribution of the $N = 130$ objects which satisfy the first criterion listed in Section 3.4.1 (i.e. a statistically acceptable redshift solution at $z_{\rm phot} \geq 6.0$) but fail to satisfy the other two criteria (see text for details).

for each candidate if Lyman α emission with a rest-frame $\rm EW_0$ in the range $0 < \rm EW_0 < 240$ Å is included as an extra free parameter in the SED-fitting procedure. This information is provided to indicate the maximum plausible redshift for each candidate. The three candidates which are listed separately at the bottom of Tables 3 and 4 pass our criteria as robust $z \geq 6$ candidates only if Lyman α emission is included in the SED-fitting procedure, and are not included in any of the subsequent analysis. The 0.6-arcsec-diameter aperture photometry for each candidate is listed in Tables B1–B3 of Appendix B, along with plots of the best-fitting SED templates and grey-scale optical/near-IR postage stamps.

4 THE UV SPECTRAL SLOPES

As discussed in the introduction, one of the most interesting, and controversial, results to emerge from the new WFC3/IR-selected LBGs has been the claim that faint LBGs ($M_{1500} \simeq -18.5$) at $z \geq 6$ display extremely blue ($\beta \simeq -3$) UV spectral slopes (e.g. Bouwens et al. 2010b; Labbé et al. 2010). Given the relatively small areas which have currently been imaged with the WFC3/IR (i.e. ≃50 square arcmin), the brightest WFC3/IR-selected $z \simeq 7$ LBGs have absolute UV luminosities of $M_{1500} \simeq -21$. It is widely agreed in the literature that at these absolute magnitudes ($L \simeq 2L^{\star}$), $z \simeq 7$ LBGs display the same UV spectral slopes ($\beta \simeq -2$) as observed for young (≃100 Myr) starbursts at redshifts $3 < z < 5$. However, in contrast, it has been claimed that the faintest LBGs at $z \simeq 7$ ($M_{1500} \simeq -18.5$) display much bluer spectral slopes; $\langle\beta\rangle = -3.0 \pm 0.2$ (Bouwens et al. 2010b).

Although this may appear to be a relatively small difference in spectral slope, it is potentially of great interest. The reason is very straightforward. While UV spectral slopes of $\beta \simeq -2$ can be comfortably reproduced by standard simple stellar population models (without recourse to ultrayoung ages or ultralow metallicities), spectral slopes of $\beta \simeq -3$ cannot, and probably require a combination of zero reddening, very young ages (i.e. ≃10–30 Myr) and a high escape fraction of photons shortwards of Lyman α (e.g. Bouwens et al. 2010b; Labbé et al. 2010). Given the potential importance of this result, not least for studies of reionization, it is clearly of interest to investigate the UV spectral slopes displayed by the sample of high-redshift LBGs derived here.

The individual values of β measured for each candidate are listed in Tables 2–4. The β values have been calculated using the following formulae:

$$\beta = 4.43(J_{125} - H_{160}) - 2.0, \quad (1)$$

$$\beta = 5.47(Y_{105} - J_{125}) - 2.0, \quad (2)$$

$$\beta = 3.91(Y_{098} - J_{125}) - 2.0, \quad (3)$$

depending on the available filters and the redshift of the candidate. To derive the above formulae, we have adopted the following pivot wavelengths for the Y_{098}, Y_{105}, J_{125} and H_{160} filters: 0.9864, 1.0552, 1.2486 and 1.5369 μm (WFC3 Instrument Handbook for Cycle 19). In order to sample as similar a range of rest-frame wavelengths as possible, and to ensure no potential contamination from Lyman α line emission, the values of β have been calculated using equation (2) or (3) for candidates at $z_{\rm phot} \leq 6.4$ and equation (1) for those candidates at $z_{\rm phot} > 6.4$. In Fig. 4, we plot the estimated UV spectral slopes versus photometric redshift for the final robust sample, split by field. Several features of this plot are worthy of comment and are briefly discussed below.

4.1 Uncertainties on derived UV spectral slopes

As can readily be seen from Fig. 4, the uncertainties on measuring β are typically large. This is simply a consequence of attempting to determine a spectral slope using two filters which are not well separated in wavelength. As an illustration, consider a galaxy at $z = 7$ with a canonical UV spectral slope of $\beta = -2$, which is detected at 5σ significance in both the J_{125} and H_{160} filters. The corresponding estimate of the UV spectral slope is $\beta = -2.0 \pm 1.3$, where the error simply reflects the photometric uncertainty. Clearly, deriving meaningful estimates of β on an individual object-by-object basis requires significantly better than 5σ photometry in both filters. One obvious method of overcoming this problem is to assume that each β measurement, although inaccurate, is at least unbiased. In which case, one can proceed to bin the data and attempt to estimate the mean value of β. However, even when adopting this approach, it is necessary to account for the wide range in β uncertainties displayed by the objects in a typical sample, by calculating a properly weighted mean:

$$\langle\beta\rangle = \frac{\sum_{i=1}^{n} \frac{\beta_i}{\sigma_i^2}}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}}, \quad (4)$$

where β_i represents an individual β measurement for a single candidate and σ_i^2 is the corresponding variance.

4.2 Average values of UV spectral slopes

The variance-weighted values of $\langle\beta\rangle$ for each subsample, and the full combined sample are listed in Table 5 where, for comparison, we also list the straight arithmetic means and standard errors. It can be seen from Table 5 that for the HUDF and ERS subsamples (and for the full combined sample), the variance-weighted mean results

Table 2. Details of the final high-redshift sample in the HUDF. The first three columns list the ID number and coordinates of each candidate. Columns 4, 5 and 6 list the photometric redshifts, the 1σ uncertainty on the photometric redshifts and the values of χ^2 for the best-fitting SED template. Column 7 lists the difference in χ^2 between the best-fitting high-redshift SED template and the alternative photometric redshift solution at low redshift (typically at $1.0 < z_{\rm phot} < 2.5$). Column 8 lists the absolute magnitude of the best-fitting SED template, where M_{1500} is measured using a 100-Å-wide filter centred on a rest-frame wavelength of 1500 Å. Column 9 lists the value of the UV spectral slope ($f_\lambda \propto \lambda^\beta$) for each candidate, derived using the formulae listed in Section 4. Column 10 lists the best-fitting photometric redshift derived including Lyman α emission as an additional free parameter, and column 11 lists the logarithm of luminosity of the corresponding best-fitting Lyman α emission line (in erg s^{-1}). Column 12 lists the results of cross-checking each candidate against existing literature studies of the WFC3/IR data in the HUDF. Matches were found with the following papers: M (McLure et al. 2010); B (Bouwens et al. 2011; robust), Bp (Bouwens et al. 2011; potential), F (Finkelstein et al. 2010), W (Wilkins et al. 2010), Y (Yan et al. 2010) and L (Lorenzoni et al. 2011).

ID	RA (J2000)	Dec. (J2000)	z	Δz	χ^2	$\Delta\chi^2$	M_{1500}	β	$z_{{\rm Ly}\alpha}$	$\log(L_{{\rm Ly}\alpha})$	Literature
HUDF_1344	03:32:36.63	−27:47:50.1	6.06	5.89–6.17	1.1	27.6	−19.3	-2.2 ± 0.7	6.16	42.2	M
HUDF_1016	03:32:35.06	−27:47:40.2	6.06	5.97–6.15	4.3	32.0	−19.6	-0.4 ± 0.6	6.33	42.4	M
HUDF_522	03:32:36.47	−27:46:41.4	6.07	5.98–6.15	4.5	151.3	−20.8	-1.9 ± 0.6	6.07	–	–
HUDF_2622	03:32:36.64	−27:47:50.2	6.11	5.95–6.38	1.3	13.4	−18.8	-2.2 ± 1.2	6.43	42.3	M
HUDF_796	03:32:37.46	−27:46:32.8	6.19	5.86–6.31	1.4	46.0	−19.9	-1.7 ± 0.6	6.50	42.7	M,F
HUDF_2836	03:32:35.05	−27:47:25.8	6.22	5.97–6.43	0.8	7.4	−18.4	-4.6 ± 2.3	6.51	42.1	M
HUDF_1692	03:32:43.03	−27:46:23.6	6.23	6.11–6.34	3.1	38.5	−19.3	-3.4 ± 0.8	6.49	42.4	M
HUDF_2743	03:32:36.52	−27:46:42.0	6.26	5.80–6.72	0.5	4.0	−18.2	-1.2 ± 2.0	6.40	41.7	M,Y
HUDF_2316	03:32:44.31	−27:46:45.2	6.31	6.03–6.54	1.2	9.1	−18.7	-1.7 ± 1.7	6.30	–	M
HUDF_2281	03:32:39.79	−27:46:33.7	6.37	6.11–6.57	0.3	5.7	−18.6	-3.2 ± 1.7	6.35	–	M
HUDF_1442	03:32:42.19	−27:46:27.8	6.37	6.17–6.55	6.4	11.2	−19.2	-1.4 ± 0.7	6.43	41.7	M,F,W,B
HUDF_2324	03:32:41.60	−27:47:04.5	6.41	6.18–6.60	0.6	5.5	−18.6	-2.9 ± 1.4	6.40	–	B
HUDF_2672	03:32:37.80	−27:47:40.4	6.45	6.14–6.67	0.5	7.9	−18.6	-2.0 ± 1.2	6.81	42.4	M
HUDF_1818	03:32:36.38	−27:47:16.3	6.57	6.35–6.72	2.1	17.3	−19.1	-2.3 ± 1.0	7.05	42.7	M,F,W,B,Y
HUDF_1473	03:32:36.77	−27:47:53.6	6.57	6.42–6.71	2.1	24.8	−19.2	-1.8 ± 0.5	6.99	42.6	M,F,W,B
HUDF_1730	03:32:43.78	−27:46:33.7	6.60	6.37–6.84	0.5	11.5	−19.1	-2.7 ± 0.8	6.59	–	M,F,W
HUDF_1632	03:32:37.44	−27:46:51.2	6.60	6.40–6.74	0.7	13.3	−19.1	-2.1 ± 0.8	6.60	–	M,F,W,B,Y
HUDF_2084	03:32:40.57	−27:46:43.6	6.61	6.39–6.80	2.6	11.9	−18.8	-2.2 ± 0.9	7.13	42.6	M,F,W,B,Y
HUDF_1995	03:32:39.58	−27:46:56.5	6.62	6.31–6.91	4.2	6.2	−18.9	-3.5 ± 0.9	6.60	–	M,F,B,Y
HUDF_658	03:32:42.56	−27:46:56.6	6.63	6.53–6.79	1.4	81.9	−20.5	-2.0 ± 0.5	6.85	42.5	M,F,W,B,Y
HUDF_2701	03:32:41.82	−27:46:11.3	6.66	6.35–6.91	2.3	4.5	−18.5	-3.3 ± 1.5	6.88	42.0	F,W,Bp,Y
HUDF_860	03:32:38.81	−27:47:07.2	6.96	6.72–7.23	1.8	31.8	−20.0	-1.4 ± 0.5	6.96	–	M,F,W,B,Y
HUDF_1102	03:32:39.55	−27:47:17.5	7.06	6.75–7.42	2.5	7.1	−19.7	-1.5 ± 0.5	7.06	–	M,F,B,Y
HUDF_1419	03:32:43.13	−27:46:28.5	7.23	6.80–7.48	5.9	9.1	−19.2	-1.3 ± 0.7	7.95	42.8	M,F,W,B,L,Y
HUDF_2641	03:32:39.73	−27:46:21.3	7.35	6.97–7.76	1.2	7.2	−18.7	-4.3 ± 1.4	8.06	42.6	M,F,B,Y
HUDF_1962	03:32:38.36	−27:46:11.9	7.36	6.80–7.73	1.0	5.5	−19.1	-3.2 ± 1.1	7.27	–	B,F,Y
HUDF_1173	03:32:44.70	−27:46:44.3	7.36	7.07–7.72	4.5	9.0	−19.8	-2.4 ± 0.6	7.36	–	M,F,B,Y
HUDF_2664	03:32:33.13	−27:46:54.5	7.45	6.98–7.89	1.9	4.2	−18.6	-4.8 ± 2.4	8.08	42.5	M,B,L
HUDF_1660	03:32:37.21	−27:48:06.2	7.52	7.24–7.76	0.9	14.4	−19.3	-2.6 ± 0.7	7.98	42.5	M,F,B,Y
HUDF_1679	03:32:42.88	−27:46:34.5	7.88	7.51–8.11	1.7	5.5	−19.1	-2.0 ± 0.8	8.80	42.7	M,F,B,L,Y
HUDF_2003	03:32:38.13	−27:45:54.0	8.49	8.08–8.75	0.9	7.5	−19.1	-2.3 ± 1.1	8.89	42.6	M,F,B,L,Y

in a significantly redder estimate of the typical value of the UV slope than the straight arithmetic mean. Interestingly, for the HUDF09-2 subsample, where the photometry is most robust (see discussion below), the difference between the two estimates is negligible.

The results listed in Table 5 indicate that the ERS subsample contains a higher percentage of objects with $\beta \leq -2$ than the other two fields (based on the variance-weighted means), although the difference is not significant. However, it is worth noting that any suggestion that the ERS candidates display bluer UV spectral slopes cannot be due to a trend for increasingly blue UV spectral slopes with decreasing UV luminosity, given that the median absolute magnitude of the ERS sample is $M_{1500} = -20.2$ compared to $M_{1500} = -19.1$ for the HUDF. Overall, our results provide no evidence that the members of the $z \geq 6.5$ LBG population display values of β significantly different from those seen in comparably luminous LBGs in the redshift interval $3 < z < 5$.

4.3 Potential for bias

It can be seen from Fig. 4 that the HUDF09-2 subsample seems to display a particularly tight distribution of UV slopes, whereas the HUDF and ERS subsamples show considerably more scatter. At least part of the explanation for this is that the HUDF09-2 subsample has the most robust WFC3/IR photometry. The reason is that, although the WFC3/IR imaging of the HUDF09-2 is deep (particularly the J_{125} data), the supporting data at other wavelengths are not, in a relative sense, as good (e.g. no B_{435} data, relatively shallow $V_{606} + i_{775}$ data, and no 4.5-μm data). As a consequence, candidates in the HUDF09-2 are required to be somewhat brighter in the near-IR in order to pass our robustness criteria (see photometry in Appendix B).

Another noteworthy point is that the bluer mean UV slope in the ERS subsample is probably connected to the relative depths of the WFC3/IR imaging in this field. Due to the fact that the J_{125} imaging in the ERS subsample is significantly deeper than the accompanying Y_{098} and H_{160} imaging, the ERS subsample is the closest of the three to being purely J_{125}-selected. It is clear that when estimating the UV spectral slope from the $J_{125} - H_{160}$ colour, selecting the sample largely on the apparent J_{125} magnitude must introduce the potential for biasing the sample towards objects with blue values of β. A proper investigation of the sources of bias, and the potential for constraining the true underlying distribution of UV spectral slopes,

Table 3. Details of the final high-redshift sample in the ERS field. Columns 1–11 list the same quantities as in Table 2. Column 12 lists the results of cross-checking each candidate against existing literature studies of the WFC3/IR data in the ERS field. Matches were found with the following papers: B (Bouwens et al. 2011; robust), Bp (Bouwens et al. 2011; potential), W (Wilkins et al. 2010) and L (Lorenzoni et al. 2011).

ID	RA (J2000)	Dec. (J2000)	z	Δz	χ^2	$\Delta\chi^2$	M_{1500}	β	$z_{Ly\alpha}$	$\log(L_{Ly\alpha})$	Literature
ERS_7086	03:32:34.75	−27:40:35.1	6.18	6.02–6.35	1.4	11.2	–20.2	-2.4 ± 0.6	6.36	42.5	–
ERS_6066	03:32:07.86	−27:42:17.8	6.19	5.87–6.41	2.7	15.8	–20.3	-2.5 ± 0.6	6.66	43.1	–
ERS_9869	03:32:15.40	−27:43:28.6	6.21	6.01–6.41	0.4	8.0	–19.8	-3.4 ± 1.0	6.44	42.5	Bp
ERS_8668	03:32:27.96	−27:41:19.0	6.22	5.88–6.55	2.0	7.6	–19.9	-2.2 ± 0.8	6.22	–	–
ERS_9100	03:32:20.24	−27:43:34.3	6.27	5.95–6.50	1.4	6.6	–19.8	-2.3 ± 0.6	6.45	42.5	Bp
ERS_7225	03:32:36.31	−27:40:15.0	6.30	6.02–6.73	4.3	11.4	–20.1	-1.7 ± 0.7	7.16	43.1	–
ERS_6438	03:32:25.28	−27:43:24.2	6.33	6.14–6.70	6.7	9.5	–20.3	-2.0 ± 0.7	7.26	43.2	W
ERS_6263	03:32:06.83	−27:44:22.2	6.36	6.14–6.59	4.9	10.0	–20.3	-2.1 ± 0.8	6.40	41.9	B
ERS_7776	03:32:03.77	−27:44:54.4	6.46	6.15–6.66	3.9	6.5	–20.0	-1.6 ± 0.8	6.58	42.3	–
ERS_5847	03:32:16.00	−27:43:01.4	6.49	6.31–6.59	2.7	17.2	–20.5	-2.4 ± 0.8	6.90	43.1	W
ERS_8987	03:32:16.01	−27:41:59.0	6.52	6.06–6.84	2.5	5.9	−19.7	-1.8 ± 1.0	6.72	42.5	B
ERS_3679	03:32:22.66	−27:43:00.7	6.55	6.42–6.71	5.5	14.5	–21.2	-1.9 ± 0.5	6.55	–	W,Bp
ERS_7412	03:32:09.85	−27:43:24.0	6.57	6.37–6.77	9.1	8.8	–20.2	-0.7 ± 0.7	7.58	43.3	–
ERS_6427	03:32:24.09	−27:42:13.9	6.65	6.37–6.88	1.2	10.3	–20.3	-2.5 ± 0.6	6.64	42.2	W,B
ERS_8858	03:32:16.19	−27:41:49.8	6.79	6.33–7.08	1.2	6.8	–20.0	-2.9 ± 0.8	6.77	–	B
ERS_7376	03:32:29.54	−27:42:04.5	6.79	6.50–6.98	0.6	5.3	–20.2	-2.4 ± 1.1	7.27	43.0	W,B
ERS_8176	03:32:23.15	−27:42:04.7	6.81	6.62–6.98	4.4	13.6	–20.1	-3.8 ± 1.1	7.73	43.3	W
ERS_7672	03:32:10.03	−27:45:24.6	6.88	6.64–7.05	3.8	7.7	–20.3	-2.4 ± 1.0	7.77	43.3	–
ERS_7475	03:32:32.81	−27:42:38.5	7.11	6.83–7.31	5.2	6.2	–20.3	-1.9 ± 0.7	7.76	43.2	–
ERS_7236	03:32:11.51	−27:45:17.1	7.18	6.99–7.35	5.0	5.4	–20.3	-3.4 ± 0.8	7.74	43.0	–
ERS_9041	03:32:23.37	−27:43:26.5	8.02	7.61–8.20	6.2	9.1	–20.0	-5.0 ± 1.5	8.16	43.0	L
ERS_10288	03:32:35.44	−27:41:32.7	8.28	7.59–8.56	2.1	7.8	–20.1	-2.3 ± 0.8	9.50	43.2	B
ERS_8584	03:32:02.99	−27:43:51.9	8.35	7.63–8.74	2.9	4.7	–20.4	-1.6 ± 1.1	9.37	43.3	B,L
ERS_8496	03:32:29.69	−27:40:49.9	6.07	5.49–6.56	4.6	9.0	−19.7	-1.8 ± 1.1	6.87	42.9	–
ERS_9923	03:32:10.06	−27:45:22.6	6.59	6.37–6.78	6.8	7.7	–20.0	-1.7 ± 1.3	7.59	43.3	–

Table 4. Details of the final high-redshift sample in the HUDF09-2. Columns 1–11 list the same quantities as in Table 2. Column 12 lists the results of cross-checking each candidate against existing literature studies of the WFC3/IR data in the HUDF09-2. Matches were found with the following papers: B (Bouwens et al. 2011; robust), Bp (Bouwens et al. 2011; potential) and W (Wilkins et al. 2010).

ID	RA (J2000)	Dec. (J2000)	z	Δz	χ^2	$\Delta\chi^2$	M_{1500}	β	$z_{Ly\alpha}$	$\log(L_{Ly\alpha})$	Literature
HUDF09-2_2459	03:33:06.30	−27:50:20.2	6.06	5.92–6.17	3.2	20.8	−19.2	-1.4 ± 1.0	6.37	42.5	–
HUDF09-2_2613	03:33:06.52	−27:50:34.6	6.08	5.90–6.22	1.0	13.2	−19.2	-3.0 ± 1.0	6.44	42.5	–
HUDF09-2_2638	03:33:06.65	−27:50:30.2	6.14	5.76–6.46	0.1	4.5	−19.0	-2.1 ± 1.1	6.44	42.2	–
HUDF09-2_1543	03:33:01.18	−27:51:22.3	6.18	6.05–6.26	0.3	22.3	−20.4	-2.3 ± 0.6	6.12	–	–
HUDF09-2_605	03:33:01.95	−27:52:03.2	6.30	6.06–6.49	0.1	6.0	−19.5	-1.9 ± 0.7	6.30	–	–
HUDF09-2_2587	03:33:04.20	−27:50:31.3	6.30	6.11–6.39	3.3	27.9	−20.3	-1.3 ± 0.6	6.28	41.8	–
HUDF09-2_1660	03:33:01.10	−27:51:16.0	6.36	6.15–6.47	3.8	7.9	−20.3	-1.1 ± 0.6	6.26	–	–
HUDF09-2_1745	03:33:01.19	−27:51:13.3	6.52	6.22–6.82	0.3	7.9	−19.4	-1.9 ± 0.6	6.98	42.7	W,B
HUDF09-2_1620	03:33:05.40	−27:51:18.8	6.61	6.30–6.93	1.7	7.0	−19.1	-1.7 ± 1.0	7.39	42.9	W,B
HUDF09-2_1721	03:33:01.17	−27:51:13.9	6.73	6.39–7.05	2.5	4.3	−19.8	-2.1 ± 0.5	6.78	42.0	–
HUDF09-2_2455	03:33:09.65	−27:50:50.8	6.82	6.73–6.89	1.9	34.4	−20.6	-1.9 ± 0.5	7.12	43.1	W,Bp
HUDF09-2_1584	03:33:03.79	−27:51:20.4	7.17	6.79–7.36	0.7	15.6	−20.6	-1.6 ± 0.5	8.03	43.3	W,B
HUDF09-2_2814	03:33:07.05	−27:50:55.5	7.30	6.90–7.66	0.5	5.7	−19.7	-2.5 ± 0.6	7.26	–	Bp
HUDF09-2_1596	03:33:03.76	−27:51:19.7	7.45	7.06–7.62	6.0	15.8	−20.4	-2.0 ± 0.5	7.95	43.1	B
HUDF09-2_2000	03:33:04.64	−27:50:53.0	7.68	7.30–7.90	1.7	15.5	−19.7	-2.2 ± 0.5	8.01	42.7	B
HUDF09-2_2765	03:33:07.58	−27:50:55.0	8.70	8.37–9.05	0.9	6.9	−20.0	-0.5 ± 0.6	8.76	41.8	B
HUDF09-2_799	03:33:09.15	−27:51:55.4	6.88	6.70–7.00	9.1	15.5	−19.5	-1.6 ± 0.6	7.67	43.1	B,W

requires detailed simulation work which, although beyond the scope of this paper, is investigated in detail by Dunlop et al. (2011).

5 STELLAR MASSES AND STAR FORMATION RATES

A key advantage of employing a template-fitting SED analysis is that stellar mass and SFR estimates can be directly derived from the best-fitting models. In this section, we use this information to investigate the relationship between stellar mass and UV luminosity, and try to determine the typical sSFR for $L^\star$ LBGs at $z \geq 6$. Throughout this section, we will repeatedly refer to the results for a subsample of 21 objects which have the most reliable SFR and stellar-mass estimates because they are detected at either 3.6 μm or 3.6+4.5 μm. In Table 6, we list the best-fitting parameters returned by our SED-fitting analysis for these objects, based on the best-fitting templates

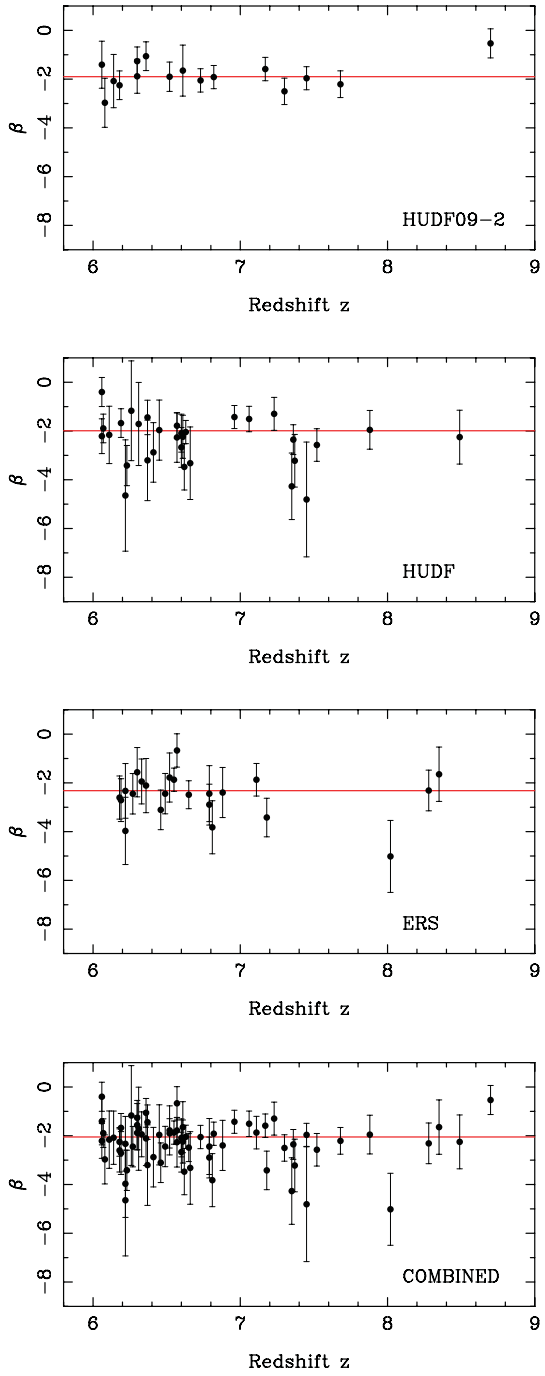

Figure 4. Plots of UV spectral slope (β) versus redshift for the final robust sample of 70 objects at $z_{\rm phot} \geq 6$. The top three panels show β versus redshift for the three separate survey fields, while the bottom panel shows β versus redshift for the full combined sample. In each panel, the horizontal line shows the variance-weighted mean value of β (see Table 5).

drawn from the full range of SFHs, metallicities and reddening described in Section 3.3. Based on photometry alone, it is very difficult to accurately constrain the SFH and metallicity of high-redshift galaxies. As a result, it is common in the literature to derive SFR and stellar-mass estimates from a much more restricted set of SED templates (typically constant-SFR, hereinafter CSF, models). In order to investigate the effect of this approach, in Table 6 we also list the best-fitting parameters derived from an SED template with a CSF, $Z = 0.2\,{\rm Z}_{\odot}$ and zero reddening (CSF model).

5.1 Stellar mass–UV luminosity relation

In Fig. 5, we plot stellar mass versus absolute UV magnitude (M_{1500}) for our final robust sample of 70 $z \geq 6$ LBGs. In the left-hand panel, we plot stellar-mass estimates based on the best-fitting SED templates drawn from the full grid of SFHs, metallicities and reddening. In the right-hand panel, we plot the stellar-mass estimates based on the CSF model alone. In both panels, the small open circles indicate those objects which are formally undetected at 3.6 μm. For these objects, the only stellar mass constraints at $\lambda_{\rm rest} \geq 4000$ Å come from the upper limits at 3.6/4.5 μm provided by the deconfusion analysis. In contrast, those objects which are detected at 3.6 μm ($\geq 2\sigma$) are plotted as large grey circles, and those objects detected at both 3.6 and 4.5 μm are plotted as the large black circles.

Based on the data presented in the left-hand panel of Fig. 5, we have used the FITXY routine (Press et al. 1992) to derive the following relationship between stellar mass and UV luminosity (L_{1500}):

$$\log\left(\frac{M_{\star}}{{\rm M}_{\odot}}\right) = (2.14 \pm 0.56)\log\left(\frac{L_{1500}}{{\rm W\,Hz^{-1}}}\right) - 37.05 \pm 4.52, \quad (5)$$

which is shown as the thick black line in the left-hand panel of Fig. 5. It is interesting to compare our equation (5) with the $M_{\star}$–$L_{\rm UV}$ relation derived by González et al. (2011) based on ≃500 B-drop galaxies at $z \simeq 4$. The $M_{\star}$–$L_{\rm UV}$ relation derived by González et al. (2011) is plotted as the thick grey line in both panels of Fig. 5, and has the form $M_{\star} \propto L_{1500}^{1.7\,\pm0.2}$. It can be seen from the left-hand panel of Fig. 5 that both relations are clearly consistent, although the $M_{\star}$–$L_{\rm UV}$ relation derived here for LBGs with a mean redshift of $z = 6.8 \pm 0.1$ is somewhat steeper.

In an earlier study, Stark et al. (2009) also explored the $M_{\star}$–$L_{\rm UV}$ relation based on $4 < z < 6$ LBGs selected from the GOODS N+S fields. At $z \simeq 4$, the data from Stark et al. (2009), based on a sample of ≃700 B-drop candidates, are entirely consistent with the $M_{\star}$–$L_{\rm UV}$ relation derived by González et al. (2011). At higher redshifts, both Stark et al. (2009) and González et al. (2011) investigated the $M_{\star}$–$L_{\rm UV}$ relation at $z \simeq 5$ and $z \simeq 6$, based on samples of V-drop and i-drop LBG candidates, respectively. Interestingly, at $z \simeq 5$, the results of both studies do appear to be consistent with a steepening of the $M_{\star}$–$L_{\rm UV}$ relation. In fact, this effect was noted by Stark et al. (2009) but, based on the available data, both authors concluded that there was no strong evidence for redshift evolution. At $z \simeq 6$, neither study had sufficient dynamic range in $L_{\rm UV}$ to constrain the slope of the $M_{\star}$–$L_{\rm UV}$ relation.

At a given UV luminosity, the range of stellar masses displayed by the LBG candidates in Fig. 5 is simply a function of their mass-to-light ratios ($M_{\star}/L_{\rm UV}$) which, in turn, are largely a function of their stellar population ages. Unfortunately, those candidates which have neither detections nor meaningful upper limits at IRAC wavelengths inevitably have stellar ages/masses which are very poorly constrained (small grey open circles in Fig. 5). These objects (which we have excluded from our determination of the best-fitting $M_{\star}$–$L_{\rm UV}$ relation) can be seen to congregate close to the lower limit which is imposed during the SED-fitting procedure by insisting that each candidate must have an age of ≥10 Myr. However, in reality, the majority of these objects can tolerate SED fits with stellar populations as old as ≃200 Myr, at which point their estimated stellar masses become an order of magnitude larger. Consequently, the apparent steepening of the $M_{\star}$–$L_{\rm UV}$ relation at faint magnitudes must be viewed with considerable caution.

Table 5. Estimates of the typical value of the UV spectral slope (β) for our final robust sample. The first two columns list the names and sizes of the different samples being considered. Columns 3 and 4 list the variance-weighted and arithmetic mean values of β, respectively, together with their corresponding uncertainties. Note that candidate HUDF09-2_2765 has been excluded from these calculations because it provides a biased estimate of β due to its high redshift ($z_{\rm phot} = 8.7 \pm 0.3$).

Sample	N	$\langle\beta_{\rm var}\rangle$	$\langle\beta_{\rm arith}\rangle$
HUDF	31	-1.99 ± 0.14	-2.40 ± 0.18
HUDF09-2	15	-1.90 ± 0.15	-1.91 ± 0.13
ERS	23	-2.32 ± 0.18	-2.52 ± 0.19
COMBINED	69	-2.05 ± 0.09	-2.33 ± 0.11

It is clear from Fig. 5 that, based on the current sample, it is not possible to determine if the $M_\star$–$L_{\rm UV}$ relation at $z \geq 6$ is steeper than at $z \simeq 4$. Indeed, our results for the 21 objects with the most reliable stellar-mass estimates are entirely consistent with the conclusion that the slope and normalization of the $M_\star$–$L_{\rm UV}$ relation do not change over the redshift interval $4.0 < z < 7.0$. However, by restricting ourselves to those objects with the most reliable stellar-mass estimates, the results presented in the left-hand panel of Fig. 5 suggest that $L^\star$ ($M_{1500} \simeq -20.2$) galaxies at $z \simeq 6.5$ have a median stellar mass of $M_\star = (2.1 \pm 1.1) \times 10^9\,{\rm M}_\odot$. Moreover, by deriving stellar-mass estimates using stellar population models covering a wide range of metallicities, SFHs and reddening, our results indicate that the full range of $M_\star/L_{\rm UV}$ displayed by $L^\star$ galaxies at this epoch could span a factor of $\simeq$50.

Within this context it is interesting to compare the left-hand and right-hand panels of Fig. 5 where the limiting effect of restricting the SED fitting to a CSF model is explored. It can immediately be seen from the right-hand panel that if we adopt the same approach as González et al. (2011) and restrict our SED fitting to the CSF model, then our stellar-mass estimates at $z \simeq 6.8$ fall into excellent agreement with the $M_\star$–$L_{\rm UV}$ relation they derived at $z \simeq 4$. Moreover, it is also clear that restricting the SED-fitting analysis to the CSF model significantly reduces (perhaps unrealistically) the scatter in the stellar-mass estimates at a given UV luminosity. Finally, as illustrated by the upper dotted line in the right-hand panel of Fig. 5, the brightest LBGs in our sample (i.e. $M_{1500} \leq 19.0$) are fully consistent with the expected $M_\star$–$L_{\rm UV}$ relation for a galaxy which

Table 6. The best-fitting parameters returned by our SED-fitting analysis of the 21 objects in our final robust sample with detections at either 3.6 or 3.6+4.5 μm. The first column lists the objects IDs. The second column lists the best-fitting SFH which is either an instantaneous burst (Burst), CSF (Const) or exponentially decaying SFR (E). For those objects where the best-fitting SFH is exponentially decaying [i.e. SFR $\propto {\rm e}^{(-t/\tau)}$], the proceeding number indicates the characteristic star-formation time-scale in Gyr (i.e. E0.2 $\Rightarrow \tau = 0.2$ Gyr). Columns 3–8 list the metallicity, age, reddening, stellar mass, SFR and χ^2 of the best-fitting SED template. For those objects where the best-fitting SFH is Burst, the SFR listed in column 7 is derived from the best-fitting model with a constant or exponentially decaying SFH. Note that the ages listed in columns 4 and 10 refer to the total age of the system, rather than a luminosity-weighted age. Columns 9–12 list the χ^2, age, stellar mass and SFR of the best-fitting CSF model with a metallicity of $0.2\,{\rm Z}_\odot$ and $A_V = 0.0$ (CSF model, see text for details). Based on this restricted set of SED parameters, those objects highlighted with a † symbol in column 9 would have been rejected because their best-fitting SED templates have an unacceptably high χ^2. The final column lists an estimate of the SFR based on the UV luminosity of each object (M_{1500}), with no correction for dust attenuation, which has been derived using the Madau, Pozzetti & Dickinson (1998) formula (corrected to a Chabrier IMF).

ID	SFH	Z ($Z_\odot$)	Age (Myr)	A_V	$M_\star$ ($10^9\,{\rm M}_\odot$)	SFR ($M_\odot\,{\rm yr}^{-1}$)	χ^2_b	χ^2_c	Age (Myr)	$M_\star$ ($10^9\,{\rm M}_\odot$)	SFR ($M_\odot\,{\rm yr}^{-1}$)	SFR$_{\rm UV}$ ($M_\odot\,{\rm yr}^{-1}$)
HUDF_1016	Burst	0.02	10	1.1	$0.9^{+0.3}_{-0.5}$	$40.0^{+13.8}_{-11.3}$	4.3	19.3†	575	$0.9^{+0.5}_{-0.4}$	$2.1^{+0.1}_{-0.1}$	3.8
HUDF_522	Burst	0.02	50	0.1	$1.8^{+0.9}_{-1.0}$	$15.0^{+6.8}_{-1.9}$	4.5	6.9	365	$1.6^{+1.0}_{-0.4}$	$6.4^{+0.3}_{-0.4}$	11.3
HUDF_658	Burst	0.50	65	0.0	$1.9^{+1.9}_{-1.0}$	$2.5^{+3.3}_{-0.3}$	1.4	5.2	725	$2.4^{+0.6}_{-0.9}$	$5.1^{+0.1}_{-0.4}$	8.6
HUDF_860	Burst	0.50	100	0.1	$2.6^{+2.6}_{-0.5}$	$1.8^{+2.0}_{-0.1}$	1.8	17.8†	645	$1.5^{+0.4}_{-0.5}$	$3.4^{+0.3}_{-0.3}$	5.4
HUDF_1173	Burst	0.02	55	0.0	$0.6^{+1.4}_{-0.5}$	$1.1^{+2.1}_{-0.4}$	4.5	4.7	455	$0.8^{+0.8}_{-0.5}$	$2.4^{+0.5}_{-0.3}$	4.5
HUDF09-2_1543	Const	0.02	725	0.1	$2.1^{+3.1}_{-1.3}$	$4.5^{+1.3}_{-2.6}$	0.3	0.4	575	$1.5^{+0.9}_{-0.3}$	$4.0^{+0.4}_{-0.4}$	7.8
HUDF09-2_2587	Const	0.02	645	0.6	$6.3^{+3.6}_{-3.4}$	$13.8^{+4.4}_{-7.6}$	3.3	9.5†	815	$2.3^{+0.6}_{-0.9}$	$4.1^{+0.1}_{-0.3}$	7.2
HUDF09-2_1660	E1.0	0.02	725	0.5	$6.9^{+1.8}_{-5.1}$	$10.0^{+2.8}_{-6.1}$	3.8	8.7†	815	$2.1^{+1.3}_{-0.5}$	$4.1^{+0.3}_{-0.3}$	7.2
HUDF09-2_1584	E0.2	0.20	725	0.1	$16.5^{+9.6}_{-13.3}$	$3.8^{+0.5}_{-3.6}$	0.7	7.6†	645	$2.3^{+0.6}_{-0.9}$	$5.3^{+0.3}_{-0.3}$	9.4
ERS_7086	Burst	0.02	25	0.1	$0.4^{+0.4}_{-0.1}$	$4.9^{+4.1}_{-2.6}$	1.4	1.7	130	$0.4^{+0.4}_{-0.6}$	$3.6^{+2.3}_{-0.5}$	6.5
ERS_6066	E0.2	1.00	645	0.1	$13.4^{+3.5}_{-8.0}$	$4.4^{+0.4}_{-0.5}$	2.7	25.4†	815	$2.4^{+0.6}_{-0.9}$	$4.4^{+0.4}_{-0.3}$	7.2
ERS_9100	E0.2	0.50	405	0.0	$1.8^{+2.6}_{-0.9}$	$2.0^{+1.3}_{-0.8}$	1.4	3.2	815	$1.4^{+0.4}_{-0.5}$	$2.6^{+0.3}_{-0.4}$	4.5
ERS_7225	E0.05	0.20	325	0.0	$6.4^{+6.4}_{-1.3}$	$0.4^{+1.9}_{-0.1}$	4.3	28.0†	815	$1.9^{+0.5}_{-0.8}$	$3.6^{+0.4}_{-0.4}$	6.0
ERS_6438	Burst	0.20	160	0.0	$4.5^{+2.6}_{-2.3}$	$3.6^{+0.4}_{-0.4}$	6.7	14.9†	725	$2.1^{+0.5}_{-0.8}$	$4.5^{+0.3}_{-0.5}$	7.2
ERS_6263	Burst	0.02	330	0.0	$0.5^{+1.1}_{-0.4}$	$2.5^{+5.3}_{-0.8}$	4.9	5.3	130	$0.4^{+0.6}_{-0.3}$	$4.3^{+4.4}_{-0.8}$	7.2
ERS_7776	E0.2	0.02	325	0.0	$1.1^{+1.1}_{-0.8}$	$2.1^{+1.1}_{-1.3}$	3.9	4.2	515	$1.0^{+0.6}_{-0.6}$	$2.9^{+0.5}_{-0.4}$	5.4
ERS_5847	E1.0	0.20	815	0.0	$3.3^{+4.9}_{-2.4}$	$4.1^{+1.1}_{-2.8}$	2.7	2.9	815	$2.4^{+2.4}_{-1.4}$	$4.4^{+0.4}_{-0.4}$	8.6
ERS_3679	Burst	0.02	725	0.2	$4.5^{+6.8}_{-2.6}$	$15.0^{+0.9}_{-10.6}$	5.5	8.3	725	$4.3^{+1.1}_{-1.5}$	$8.9^{+0.3}_{-0.8}$	16.4
ERS_7412	Burst	1.00	10	0.2	$0.4^{+0.8}_{-0.3}$	$11.4^{+7.1}_{-2.9}$	9.1	12.2†	255	$0.8^{+0.8}_{-0.5}$	$3.9^{+1.1}_{-0.5}$	6.5
ERS_6427	E0.2	0.02	575	0.0	$4.4^{+4.4}_{-1.6}$	$2.1^{+0.4}_{-0.8}$	1.2	6.6	725	$2.0^{+0.5}_{-0.4}$	$4.1^{+0.4}_{-0.5}$	7.2
ERS_7376	Const	0.20	455	0.0	$1.1^{+1.6}_{-0.9}$	$3.5^{+1.1}_{-2.5}$	0.6	0.6	455	$1.1^{+1.1}_{-0.8}$	$3.5^{+0.9}_{-0.5}$	6.5

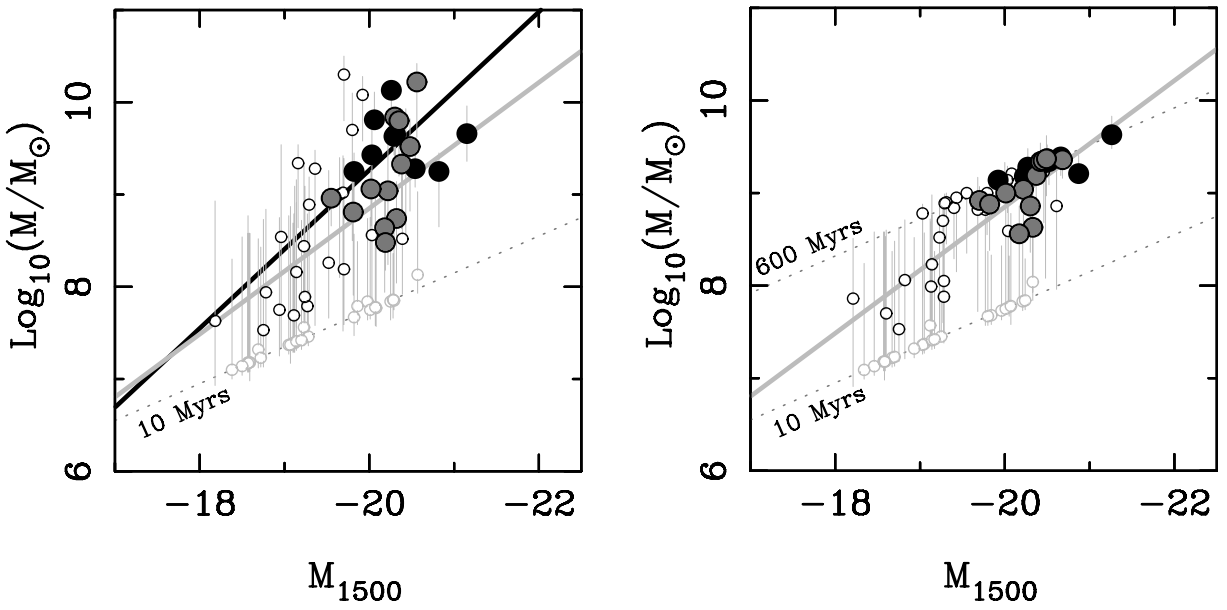

Figure 5. Stellar mass versus absolute UV magnitude (M_{1500}) for the final robust sample of 70 $z \geq 6$ LBGs, where M_{1500} is derived from the best-fitting SED models using a 100-Å-wide filter centred on a rest-frame wavelength of 1500 Å. In each panel, the small open circles are objects with non-detections (i.e. $\leq 2\sigma$) at 3.6 μm, the large grey circles are objects which are detected at 3.6 μm and the large black circles are objects detected at both 3.6 and 4.5 μm. In the left-hand panel, the stellar-mass measurements are based on our SED-fitting analysis using the full range of SFHs, metallicities and reddening described in Section 3.3. In the right-hand panel, the stellar-mass measurements are based on a single set of SED templates with a CSF, $Z = 0.2\,Z_{\odot}$ and zero reddening (CSF model, see text for details). In the left-hand panel, the thick black line is our best-fitting $M_{\star}$–$L_{\rm UV}$ relation (equation 5) and the dotted line indicates the lower limit enforced by insisting that each candidate has an age $\geq$10 Myr. Those objects which lie on the lower limit (small grey open circles) have very poorly constrained stellar masses and were not included in the derivation of the best-fitting $M_{\star}$–$L_{\rm UV}$ relation. In both panels, the thick grey line is the $M_{\star}$–$L_{\rm UV}$ relation (corrected to a Chabrier IMF) derived by González et al. (2011) using a large sample of LBGs at $z \simeq 4$. In the right-hand panel, the upper dotted line indicates the expected $M_{\star}$–$L_{\rm UV}$ relation for a CSF model which has been forming stars for $\simeq$600 Myr (i.e. since $z \simeq 20$ for an object at the mean redshift of the final robust sample; $z = 6.8 \pm 0.1$).

has been forming stars at a constant rate for $\simeq$600 Myr. Importantly, at the mean redshift of the final robust sample ($z = 6.8 \pm 0.1$), 600 Myr represents $\simeq$80 per cent of the age of the Universe. Indeed, the primary cause of the clustering of objects around the $M_{\star}$–$L_{\rm UV}$ relation corresponding to $\simeq$600 Myr of constant star formation is the requirement imposed during the SED fitting that objects must be younger than the age of the Universe. The underlying cause is simply that (with no dust reddening) the CSF models are bluer than the observed photometry unless their age is close to the maximum allowable at this epoch. In summary, although the results shown in both panels of Fig. 5 are broadly compatible, it is clear that adopting a restricted set of SED templates may well provide a misleadingly low estimate of the true level of scatter in stellar mass at a given UV luminosity.

Before moving on to consider the relationship between stellar mass and SFR, it is worth remembering that one of the principal motivations for studying the $M_{\star}$–$L_{\rm UV}$ relation at high redshift is to constrain the galaxy stellar-mass function (e.g. McLure et al. 2009). The results presented in Fig. 5 clearly illustrate that in order to successfully constrain the stellar-mass function at $z \geq 6.5$ it will be necessary to constrain the $M_{\star}$–$L_{\rm UV}$ relation at UV luminosities substantially fainter than $L^{\star}$. Over the next 3 years, the new Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS; co-PIs: S. Faber and H. Ferguson; see Grogin et al. 2011 and Koekemoer et al. 2011) offers the prospect of significant progress. The deep portion of the CANDELS will provide Y_{105}, J_{125} and H_{160} WFC3/IR imaging to $m_{\rm AB} \simeq 28(5\sigma)$ over an area of $\simeq$150 square arcmin in the GOODS N+S fields. The CANDELS programme should therefore provide a sample of $\gtrsim$200 robust $z \simeq 7$ candidates in the magnitude range $-19 > M_{1500} > -20$, all covered by the deep IRAC imaging available in the GOODS N+S. A sample of this size should be sufficient to obtain robust constraints on the typical $M_{\star}/L_{\rm UV}$ at $M_{1500} \simeq -19$ by employing a stacking analysis to provide the necessary IRAC photometry. Obtaining constraints on the $M_{\star}$–$L_{\rm UV}$ relation at magnitudes as faint as $M_{1500} \simeq -18.5$ (i.e. $\simeq 0.2L^{\star}$ at $z \simeq 7$) will rely on stacking the final epoch 2 WFC3/IR imaging of the HUDF into the forthcoming, ultradeep, IRAC data being obtained as part of the Cycle 7 *Spitzer* warm mission GO-70145 (PI: Labbé).

5.2 Star formation rate versus stellar mass

Over recent years, it has become clear that studying the ratio of the current SFR to the previously assembled stellar mass (SFR/$M_{\star}$), the so-called sSFR, can provide a useful insight into the average SFH of a galaxy population. Studies of star-forming galaxies in the SDSS at $z \simeq 0.1$ (Brinchmann et al. 2004), in the Extended Groth Strip at $z \simeq 1.0$ (Noeske et al. 2007) and in the GOODS fields at $z \simeq 1.0$ and $z \simeq 2.0$ (Daddi et al. 2007; Elbaz et al. 2007) have consistently shown a correlation of the form SFR $\propto M_{\star}^{0.9}$ (sSFR $\propto M_{\star}^{-0.1}$), over a wide dynamic range in stellar mass. Moreover, these studies have shown that compared to a median value of sSFR $\simeq$2.5 Gyr^{-1} at $z = 2$ (Daddi et al. 2007), the normalization of the sSFR–$M_{\star}$ relation has decreased by a factor of $\simeq$40 over the last 10 Gyr.

Indeed, there is some evidence that a sSFR of $\simeq$2.5 Gyr^{-1} may represent the maximum sustainable SFR for galaxies at $z \geq 2$. A recent study by Karim et al. (2011) found that high-mass ($M_{\star} \geq 10^{10}\,M_{\odot}$) star-forming galaxies in the COSMOS field consistently

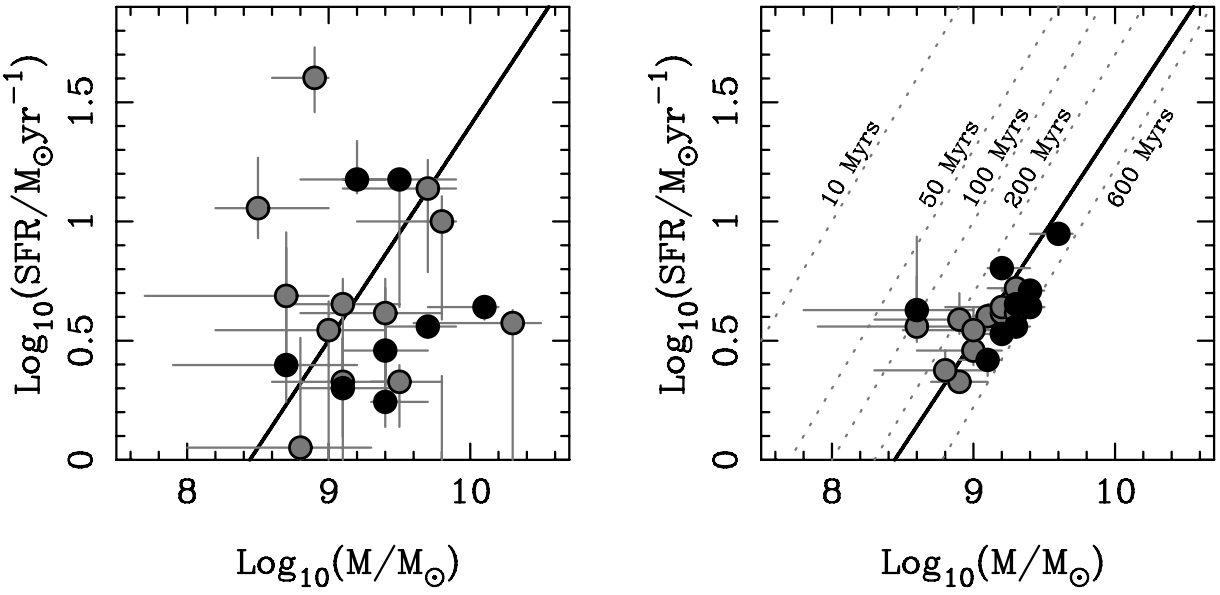

Figure 6. SFR versus stellar mass for the 21 objects in our final robust sample with IRAC detections at either 3.6 μm (grey) or 3.6+4.5 μm (black). In the left-hand panel, the SFRs and stellar masses have been measured from the best-fitting SED template drawn from the full range of SFHs, metallicities and reddening described in Section 3.3. In the right-hand panel, the SFRs and stellar masses have been estimated from the best-fitting CSF model (see text for details). The 1σ errors on both parameters have been calculated by determining the $\Delta\chi^2 = 1$ interval, after marginalization over all other free parameters. The solid line in both panels is the SFR–$M_\star$ relation derived by Daddi et al. (2007) for star-forming galaxies at $z \simeq 2$ and corresponds to a sSFR of $\simeq$2.5 Gyr^{-1}. In the right-hand panel, the dotted lines illustrate how the SFR–$M_\star$ relation for a galaxy with a CSF and zero reddening varies as a function of stellar population age.

follow a steeper sSFR–$M_\star$ relation than determined by previous studies (sSFR $\propto M_\star^{-0.4}$). However, by $z \simeq 3$, the results of Karim et al. (2011) suggest that the sSFR–$M_\star$ relation flattens at masses of $M_\star \leq 10\,\mathrm{M}_\odot$, becoming consistent with sSFR $\propto M_\star^{-0.1}$. Karim et al. (2011) interpret this flattening as a result of a natural limit to the sustainable sSFR of $\simeq$2.5 Gyr^{-1}, corresponding roughly to the inverse of the typical galaxy free-fall time-scale. Within this context, it is interesting to note that González et al. (2010) recently found that the median sSFR of a sample of 12 z-drop candidates at $z \simeq 7$ was also $\simeq$2.5 Gyr^{-1}.

In Fig. 6, we plot SFR versus stellar mass for the 21 objects which have IRAC detections and therefore the most robust SFR and stellar-mass estimates. In the left-hand panel, the stellar mass and dust-corrected SFR estimates have been taken from the best-fitting SED templates drawn from the full range of SFHs, metallicities and dust reddening described in Section 3.3. In contrast, in the right-hand panel, the stellar mass and SFR estimates have been taken from the best-fitting CSF model. Although the distribution of objects in the two panels is significantly different, both provide a consistent estimate for the typical sSFR. In the left-hand panel, the median sSFR is 1.9 ± 0.8 Gyr^{-1}, while in the right-hand panel, the median sSFR is 2.6 ± 0.4 Gyr^{-1}. Both estimates are clearly consistent with the typical sSFR value for $z \simeq 2$ star-forming galaxies estimated by Daddi et al. (2007). To illustrate this point, in both panels of Fig. 6, the thick solid line is the best-fitting SFR–$M_\star$ relation from Daddi et al. (2007) which corresponds to a sSFR of $\simeq$2.5 Gyr^{-1}.

Consequently, taken at face value, our results provide additional support to the conclusion that a direct proportionality between SFR and stellar mass is still viable at $z \simeq 6.5$, and that the corresponding sSFR of $\simeq$2.5 Gyr^{-1} may correspond to a physical limit on the maximum sustainable SFR. However, it is clear from the left-hand panel that allowing a reasonable range of SFHs, metallicities and dust reddening leads to a large scatter in the SFR at a given stellar mass. Consequently, although the data shown in the left-hand panel are consistent with SFR and stellar mass being roughly proportional, they are also entirely consistent with star formation and stellar mass being entirely unrelated. In contrast, the results shown in the right-hand panel suggest that SFR and stellar mass are well correlated, lying along a SFR–$M_\star$ relation with a slope close to unity and a normalization consistent with a sSFR of $\simeq$2.5 Gyr^{-1}.

However, it is worth noting that the apparently simple picture presented in the right-hand panel of Fig. 6 probably reflects limitations of relying on the CSF model, rather than offering genuine physical insight into high-redshift star formation. The simple reason for this caution is that the agreement is largely inevitable when you only consider SEDs with constant star formation and no reddening. In this situation, each object is required to lie on a relation with a slope of unity, with its position on the SFR–$M_\star$ plane simply determined by the best-fitting age. To illustrate this point, we have plotted the expected SFR–$M_\star$ relations for CSF models of various stellar population ages as the dotted lines in the right-hand panel of Fig. 6. This demonstrates that, provided the typical stellar population age lies in the range 200–600 Myr, the resulting SFR–$M_\star$ relation will automatically have a slope close to unity, and result in a typical sSFR consistent with $\simeq$2.5 Gyr^{-1}.

In summary, although it is possible to constrain the typical sSFR of $L^\star$ LBGs at $z \simeq 6.5$, the limitations of the current sample do not allow meaningful constraints to be placed on the form of the SFR–$M_\star$ relation. In order to resolve this issue, it will be necessary to obtain much larger samples of $z \geq 6$ LBGs with stellar masses $M_\star \geq 10^{8.5}\,\mathrm{M}_\odot$. Within this context, the new CANDELS WFC3/IR imaging data should prove decisive. The wide portion of the CANDELS will prove $J_{125} + H_{160}$ imaging to a depth of $m_{\mathrm{AB}} \simeq 27(5\sigma)$ over an area of $\simeq$0.2 deg^2, all of which is covered by deep IRAC imaging at $3.6 + 4.5$ μm ($m_{\mathrm{AB}} \simeq 26$, 5σ) provided by the Spitzer Extended Deep Survey (SEDS; PI: G. Fazio). The combination of

the CANDELS and SEDS should therefore allow the SFR–$M_\star$ relation at $z \simeq 6.5$ to be investigated using a sample of $\gtrsim$250 LBGs with reliable stellar-mass estimates of $M_\star \geq 10^{8.5}\,\mathrm{M}_\odot$.

5.3 The effect of nebular emission

Recent work has suggested that nebular continua and line emission might contribute to the observed SEDs of $z \sim 6$–7 galaxies (e.g. Ono et al. 2010; Schaerer & de Barros 2010). As discussed by Robertson et al. (2010), galaxies with strong UV continua can typically be fitted using pure stellar populations with ages of a few hundred Myr, or by much younger populations ($\leq$few Myr) with significant nebular contributions and an implied low escape fraction ($f_{\rm esc}$) of Lyman continuum photons. Such nebular solutions can yield much lower stellar masses than those in the purely stellar case (Ono et al. 2010).

In order to quantify this degeneracy and its possible effect on our derived physical properties, we have examined in more detail the subsample of 21 galaxies detected in the 3.6-μm IRAC band (nine of which are also detected at 4.5 μm). For these objects, it is possible to investigate whether the IRAC detections can provide a valuable discriminant between the nebular and stellar solutions, given the location of prominent nebular lines, such as Hβ and [O III] 5007 Å, at the redshifts of interest.

5.3.1 *Nebular emission methodology*

As before, we use the BC03 models to generate a set of spectral templates based on a Chabrier IMF. For the models presented here, we use a representative exponentially decaying SFH ($\tau \simeq 0.4$ Gyr) and, motivated by the fact that high-redshift galaxies often exhibit low metallicities (e.g. Finkelstein et al. 2011), we consider both solar ($Z = \mathrm{Z}_\odot$) and one-fifth solar ($Z = 0.2\,\mathrm{Z}_\odot$) models. In addition to exponentially decaying SFHs, we also investigated models with constant star formation, but found that these did not significantly alter our results. The contribution of the nebular continuum and line emission is computed in the manner of Robertson et al. (2010), providing nebular emission models similar to those calculated by Ono et al. (2010). The strength of the nebular emission is tied to the number of ionizing photons per second ($N_{\rm Lyc}$), calculated from the stellar population model via the Hβ luminosity (in erg s^{-1}):

$$L(\mathrm{H}\beta) = 4.78 \times 10^{-13}(1 - f_{\rm esc})N_{\rm Lyc}. \quad (6)$$

Other H I line intensities follow from ratios predicted by standard recombination theory (Osterbrock & Ferland 2006). Lines from common metallic species are included using relative intensities given by Anders & Fritze-v. Alvensleben (2003) assuming the gas-phase metallicity is either $\mathrm{Z}_\odot$ or 0.2 $\mathrm{Z}_\odot$. We use the method of Brown & Mathews (1970) to calculate the strength of bound–free and free–free continuum emission, and use results from Osterbrock & Ferland (2006) for the two photon emission from H.

When fitting the SED models, we consider two fixed values of the escape fraction, $f_{\rm esc} = 0.2$ (stellar and nebular emission) and $f_{\rm esc} = 1$ (purely stellar emission). The value of $f_{\rm esc} = 0.2$ is motivated by direct observations of the Lyman continuum in galaxies at $z \sim 3$ (Shapley et al. 2006) and typical values of $f_{\rm esc}$ required for star-forming galaxies to maintain reionization at $z \sim 7$ (Robertson et al. 2010). As we are primarily interested in how nebular emission might alter the inferred stellar mass and age, we do not include the possible effects of Lyman α emission or reddening. An example SED fit featuring nebular emission is shown in Fig. 7 and the 0.2-$\mathrm{Z}_\odot$ nebular SED fits for all 21 objects can be found in Appendix B.

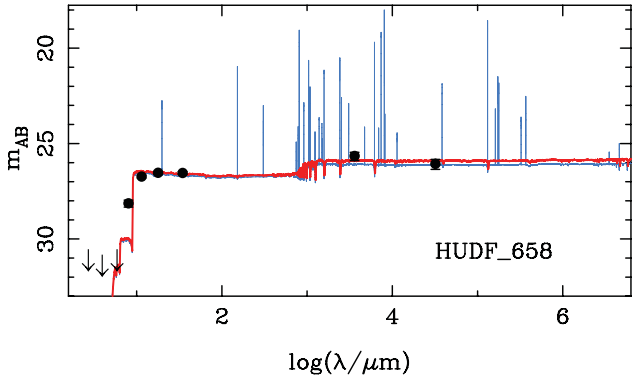

Figure 7. An example of our SED-fitting procedure using models which incorporate nebular continua and line emission. In this illustration, the pure stellar continuum model ($f_{\rm esc} = 1$) is shown as the red (thick) line and the template featuring nebular emission ($f_{\rm esc} = 0.2$) is shown as the blue (thin) line. For clarity, the galaxy template featuring nebular emission has been displaced by 0.1 mag (both models have a metallicity of $Z = 0.2\,\mathrm{Z}_\odot$).

5.3.2 *Nebular emission results*

For the models with subsolar metallicity, it is found that, compared to the purely stellar templates ($f_{\rm esc} = 1$), the templates which include nebular emission ($f_{\rm esc} = 0.2$) provide a better fit to 14 out of the 21 objects, and all nine of the objects with detections in both IRAC bands. The data favour the nebular emission models because of the predominantly blue $[3.6] - [4.5]$ μm observed colours that are easily reproduced by including rest-frame optical line emission. In contrast, the shape of the purely stellar model SEDs redwards of the Balmer/4000 Å break cannot easily accommodate these blue rest-frame optical colours.

For the majority of this subsample (16 out of 21), the best-fitting ages are still $\geq$100 Myr with stellar masses reduced by less than a factor of 2.5 (see Fig. 8). In contrast, the best-fitting models to the remaining five galaxies have significantly lower stellar masses (by more than a factor of 10 in two cases). Interestingly, all five

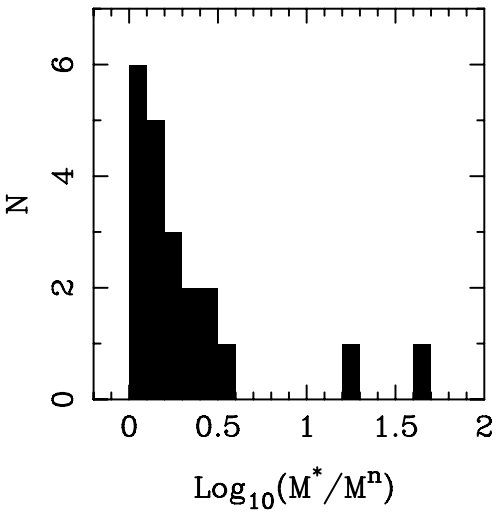

Figure 8. A histogram of the ratio of the best-fitting stellar masses returned by fitting $Z = 0.2\,\mathrm{Z}_\odot$ SED models with ($M^{\rm n}$) and without ($M^\star$) nebular continua and line emission for the 21 objects with 3.6-μm IRAC detections. For 16 out of the 21 objects, the stellar-mass estimates differ by less than a factor of 2.5. However, it can be seen that there are two objects for which the nebular fits return stellar masses which are an order of magnitude smaller than the stellar masses returned by fits without nebular emission.

of these objects are detected only in the 3.6-μm band and have best-fitting ages which are perhaps unphysical (median age 9 Myr). Using solar metallicities, we find eight galaxies with significantly altered parameters, including all five found for models with $Z = 0.2\,Z_{\odot}$.

In summary, with the current precision of the *HST* and IRAC photometry, we are unable to draw firm conclusions about the possible presence of nebular emission in these sources. However, given that the inferred stellar masses from the models including nebular emission are generally similar to those inferred from purely stellar models, we conclude that the results presented in Sections 5.1 and 5.2 appear to be robust to the inclusion of $f_{\rm esc} = 0.2$ nebular emission, especially for those objects with detection in both IRAC bands.

6 COMPARISON WITH PREVIOUS STUDIES

As discussed in the introduction, the availability of the various new WFC3/IR data sets has led to a proliferation of papers focused on $z \geq 6$ LBGs. With authors each applying their individual candidate selection procedures, and in many cases using their own independent reductions of the publicly available data, it is difficult to obtain a clear overview of the subject and to identify whether different studies are in good agreement or not. Consequently, in this section, we compare our final robust sample with those derived elsewhere in the literature (on a field-by-field basis), highlighting the objects we have in common and investigating the properties of previously published high-redshift candidates which are not included in our final robust sample.

Given that one of the primary motivations for this study was to derive a sample of high-redshift candidates which is as robust as possible, and that previous samples of WFC3/IR high-redshift candidates were selected for a variety of purposes, it is not the case that we regard any object not included in our final sample as a low-redshift interloper. In fact, as the proceeding discussion will demonstrate, each of the candidates from the literature samples falls into one of four categories. The first category consists of objects which are in common with our final sample of 70 $z \geq 6$ LBGs and we therefore regard them as being robust. The second category consists of objects which were not included in our final sample (because they failed to meet one or more of our adopted criteria), but which nevertheless our analysis suggests are likely to be at high redshift. The third category consists of objects which our analysis suggests are likely to be at low redshift, but do have an acceptable (albeit lower probability) solution at high redshift. The fourth category consists of those objects which our analysis suggests are very unlikely to be at high redshift. Throughout the discussion in this section, we have attempted to make it as clear as possible which category each of the candidates falls into. Finally, it should be noted that where a research group has published a number of studies of a particular survey field, we only discuss the results from the most recent study, under the assumption that they supersede any previous work.

6.1 HUDF

6.1.1 McLure et al. (2010)

In McLure et al. (2010), we published our initial analysis of the HUDF WFC3/IR data set, providing a list of $N = 49$ high-redshift candidates with $z_{\rm phot} \geq 5.9$. The candidate selection procedure employed in McLure et al. (2010) was broadly similar to that adopted here, with the most noteworthy difference between the two analyses being that in this work we have directly employed deconfused IRAC photometry in the candidate-selection procedure.

Of the $N = 31$ objects identified in the final HUDF sample listed in Table 2, $N = 28$ are in common with the sample derived in McLure et al. (2010), demonstrating an excellent level of agreement between the two studies. However, there are $N = 21$ candidates published in McLure et al. (2010) which do not feature in the final robust sample derived here. The reason behind this is that the primary aim of McLure et al. (2010) was to provide an estimate of the $z = 7$ and 8 galaxy luminosity functions. Consequently, the McLure et al. (2010) sample was designed to be as *complete* as possible, and therefore contained all *potential* $z \geq 5.9$ LBGs revealed by the SED-fitting analysis, irrespective of whether or not they also displayed an acceptable low-redshift solution.[5] In contrast, the principal aim of this study is to derive a sample of $z \geq 6$ LBG candidates which is as robust as possible, which means, in effect, requiring that any alternative low-redshift solutions can be statistically excluded. Although all $N = 21$ of the additional candidates listed in McLure et al. (2010) also feature in the initial catalogues derived here, all of them were excluded from our final robust HUDF sample because the best-fitting alternative solutions at low redshift could not be excluded at the $\Delta\chi^2 \geq 4$ (95 per cent) level.

6.1.2 Bouwens et al. (2011)

Based on their analysis of the HUDF data set, Bouwens et al. (2011) list a total of $N = 31$ robust high-redshift candidates, which are a mixture of z-drops and Y-drops. Of these $N = 31$ candidates, 18 also featured in our final robust sample of HUDF candidates listed in Table 2. However, it is clearly of interest to investigate why the remaining 13 objects identified by Bouwens et al. (2011) do not feature in our final robust sample.

First, we should note that six of the 13 additional objects (UDFz-38537518, UDFy-37588003, UDFy-33446598, UDFy-39347255, UDFy-40338026 and UDFy-42406550) are simply too faint to make it into our final robust sample. None of these six objects is bright enough (in a 0.6-arcsec-diameter aperture) to provide a $\geq 5\sigma$ detection in any of the WFC3/IR bands, and all six are fainter than any of our robust HUDF candidates. Consequently, this leaves a total of seven high-redshift candidates listed by Bouwens et al. (2011) which could, in principle, also feature in our final robust sample.

Our SED-fitting analysis suggests that two of the additional objects (UDFz-44746449 and UDFy-43086276) are likely to be at high redshift ($z_{\rm phot} = 8.1$ and 8.3, respectively), but just failed to make it through to our final robust sample because the competing low-redshift solutions could not be ruled out at $\geq$95 per cent confidence. Further two additional objects (UDFz-42567314 and UDFz-42247087) also have primary photometric redshift solutions at $z_{\rm phot} \geq 6.0$, but were subsequently rejected because they were either too close to the WFC3/IR array edge (UDFz-42567314[6]) or deemed to have unreliable photometry due to contamination from a nearby, bright, low-redshift galaxy (UDFz-42247087). Of the remaining three objects, one (UDFy-37796001) does have an acceptable solution at $z_{\rm phot} \geq 8$, but was rejected because our analysis suggests that the alternative solution at $z_{\rm phot} \simeq 2$ is marginally

[5] Note that the alternative low-redshift solutions were also listed by McLure et al. (2010).

[6] Reported as ID = 1144 in McLure et al. (2010).

preferred. The other two (UDFz-37296175 and UDFy-37636015) were rejected because our SED-fitting analysis returned a primary photometric redshift solution at $z_{\rm phot} \simeq 5$.

Finally, it can be seen from Table 2 that our final robust sample contains 13 objects which are not featured in the Bouwens et al. (2011) robust candidate list. However, the noteworthy feature of these objects is that the vast majority (11/13) are at $z_{\rm phot} \leq 6.5$, whereas the Bouwens et al. colour–colour, selection criteria are tuned to select objects at $z_{\rm phot} \geq 6.5$. The two exceptions (HUDF_1730 and HUDF_2701) have been identified by several different studies (see Table 2 for details) and one (HUDF_2701) does feature in the Bouwens et al. (2011) list of potential high-redshift candidates.

6.1.3 Finkelstein et al. (2010)

In their analysis of the WFC3/IR HUDF data set, Finkelstein et al. (2010) used a similar template-fitting technique to that employed in both McLure et al. (2010) and this work, and used each candidate's photometric redshift probability density function in the construction of their final list of $N = 31$ candidates at $6.3 < z_{\rm phot} < 8.6$. As part of their analysis, Finkelstein et al. (2010) conducted a detailed comparison between their final list of high-redshift candidates and the McLure et al. (2010) sample, finding a good level of agreement between the two studies.

As might be expected, the overall agreement between the analysis of Finkelstein et al. (2010) and the final robust HUDF sample derived here is still good. In the redshift range covered by both studies, our final robust HUDF sample consists of $N = 22$ candidates at $z_{\rm phot} > 6.3$, 18 of which are in common with Finkelstein et al. (2010). The four additional candidates which feature in our final robust sample are HUDF_2281, HUDF_2324, HUDF_2672 and HUDF_2664 (see Appendix B for plots of the SED fits).

Of the $N = 31$ candidates in the Finkelstein et al. (2010) sample, $N = 19$ also feature in the final HUDF sample derived here. However, this still leaves a total of 12 candidates from Finkelstein et al. (2010) which do not feature in our final sample. All 12 of these additional candidates do feature in our original HUDF catalogues, but were excluded from the final robust sample for a number of different reasons. One object (FID 3022) was excluded from our sample because it is too faint ($J_{125} \geq 29$) to provide a robust high-redshift solution, and further four objects (FIDs 640, 1818, 2013 and 2432) were excluded because they were judged to have photometry which was potentially contaminated by bright, nearby, low-redshift galaxies. For the remaining seven objects (FIDs 200, 213, 567, 653, 1110, 1566 and 2055), our SED-fitting analysis does indicate that the primary photometric redshift solution is at $z_{\rm phot} \geq 6.3$. However, all seven objects were excluded from the final robust sample because our analysis suggested that the alternative low-redshift solution could not be ruled out with $\geq$95 per cent confidence.

6.1.4 Yan et al. (2010)

In their analysis of the HUDF, Yan et al. (2010) used z-drop and Y-drop criteria to identify a sample of $N = 35$ high-redshift candidates at $z \simeq 7$ and $z \simeq 8$. Excluding a likely transient, Yan et al. (2010) list a total of 20 z-drop candidates, 14 of which are in common with our final robust HUDF sample. Of the six z-drops listed by Yan et al. (2010), which do not make it into our final robust HUDF sample, two (A046 and A056) were excluded because their alternative low-redshift solutions could not be ruled out with $\geq$95 per cent confidence, one (A017) was excluded because its photometry was contaminated by a bright, low-redshift, galaxy, and one (A008) was rejected because it lies too close to the array edge. The final two z-drops (A055 and A062) listed by Yan et al. (2010) do not feature in any of our catalogues and do not appear to be robust objects based on our reduction of the epoch 1 HUDF data set.

Yan et al. (2010) list a total of 15 Y-drop candidates in the HUDF. Of these 15 candidates, only two (B092 and B115) make it through to our final robust sample. Of the 13 Y-drops listed by Yan et al. (2010), which do not feature in our final sample, our analysis suggests that five (B041, B088, B114, B117 and SB27) do have acceptable high-redshift photometric redshift solutions, but were excluded because they all have alternative low-redshift solutions which cannot be ruled out at the $\geq$95 per cent confidence level. Further two objects (B087 and B094) also feature in our original catalogues but, based on our 0.6-arcsec-diameter photometry, are not Y-drops and have primary photometric redshift solutions at $z_{\rm phot} \leq 7$. The remaining six candidates (SB30, SD02, SD05, SD15, SD24 and SD52) do not appear as robust objects in our reduction of the epoch 1 HUDF data set. Finally, we note that Yan et al. (2010) also identify a sample of 23 J-drops in the HUDF, none of which features in our final robust HUDF sample.

6.1.5 Wilkins et al. (2010)

Wilkins et al. (2010) identify a total of 11 z-drop candidates in the HUDF, nine of which also feature in our final robust sample. Of the two additional candidates listed by Wilkins et al., our analysis suggests that one (HUDF.z.6497) does have an acceptable solution at $z_{\rm phot} \geq 6$, but was excluded because the primary photometric redshift solution lies at $z_{\rm phot} = 3.5$. The other object (HUDF.z.6433[6]) was rejected because it lies close to the array edge and was therefore deemed to have unreliable photometry.

6.1.6 Lorenzoni et al. (2011)

Based on their analysis of the HUDF data set, Lorenzoni et al. (2011) identify a sample of six Y-drop candidates. Of these six candidates, three (HUDF.YD1, HUDF.YD3 and HUDF.YD4) make it into our final robust HUDF sample. Of the remaining three candidates, our SED-fitting analysis suggests that two (HUDF.YD2 and HUDF.YD8) have an acceptable $z \simeq 8$ photometric redshift solution, but were excluded from our final robust sample because they both have an alternative low-redshift solution which cannot be securely ruled out (i.e. $\Delta\chi^2 \leq 4$). The remaining candidate (HUDF.YD9) does not appear as a robust object in any of our catalogues.

6.2 ERS

6.2.1 Bouwens et al. (2011)

The robust ERS sample derived by Bouwens et al. (2011) consists of $N = 19$ objects in total, 13 z-drops at $z \simeq 7$ and six Y-drops at $z \simeq 8$. Of the 13 z-drops listed by Bouwens et al. (2011), only five appear in our final robust sample (see Table 3). Of the remaining eight additional z-drops listed by Bouwens et al. (2011), one object (ERSz-2352941047) is too faint ($J_{125} \geq 27.5$) to produce a robust high-redshift solution based on our criteria, leaving seven additional z-drops to account for. Of these, three (ERSz-2150242362, ERSz-2225141173 and ERSz-2354442550) have statistically acceptable photometric redshift solutions at $z_{\rm phot} \geq 6.5$, but were excluded from the final robust sample because it was not possible to rule out the

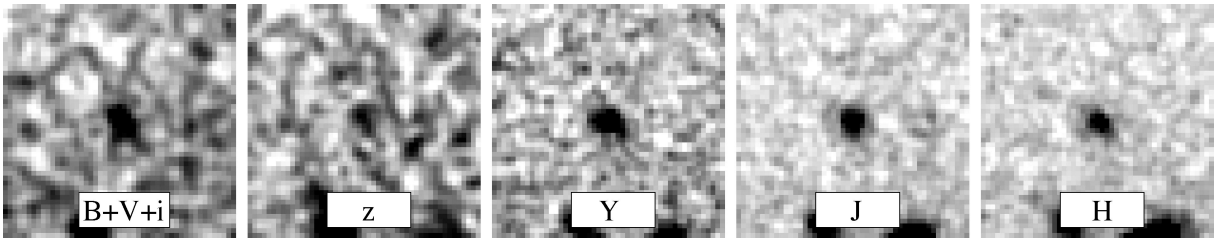

Figure 9. Stacked postage-stamp images of three $z \simeq 7$ candidates (ERS.z.80252, ERS.z.47667 and ERS.z.20851) published by Wilkins et al. (2010). All three of these objects were rejected as high-redshift candidates by our SED-fitting analysis, principally due to low signal-to-noise ratio detections in the blue optical bands. This figure demonstrates that although there is a drop in flux between the z_{850} and Y_{098} filters, these objects are also clearly detected in a stack of the $B_{435} + V_{606} + i_{775}$ GOODS imaging, suggesting that they are unlikely to be at $z \simeq 7$.

alternative low-redshift solutions at ≥95 per cent confidence. One further object (ERSz-2150943417) was rejected because based on our photometry it was not possible to obtain a statistically acceptable solution at high redshift. Of the final three objects, two (ERSz-2111644168 and ERSz-2432842478) have acceptable solutions at $z_{\rm phot} \geq 6$ but were excluded because our primary photometric redshift solution lies at $z_{\rm phot} \leq 2$. The final object (ERSz-2056344288) does not have an acceptable high-redshift solution based on our 0.6-arcsec-diameter aperture photometry.

Of the six *Y*-drops listed by Bouwens et al. (2011), two (ERSY-2354441327 and ERSY-2029843519) make it into our final robust sample. Of the four additional *Y*-drops listed by Bouwens et al. (2011), one object (ERSY-2377942344) is too faint in a 0.6-arcsec-diameter aperture ($J_{125} \geq 27.5$) to produce a robust high-redshift solution based on our criteria, leaving three additional *Y*-drops to be accounted for. Of these three objects, two (ERSY-2399642019 and ERSY-2251641574) have acceptable primary photometric redshift solutions at $z \geq 7.5$, but were excluded because the alternative low-redshift solutions could not be ruled out. The final object, ERSY-2306143041,[7] was rejected because our primary photometric redshift solution is at $z_{\rm phot} \leq 2$.

6.2.2 Lorenzoni et al. (2011)

Lorenzoni et al. (2011) identify a total of nine *Y*-drop candidates in the ERS field (five of which, marked with *, are described as 'more marginal candidates'). Of these nine candidates, only two (ERS.YD1 and ERS.YD2*) make it into our final robust sample. Of the remaining seven candidates, our analysis suggests that three (ERS.YD5*, ERS.YD6 and ERS.YD9*) have an acceptable solution at $z_{\rm phot} > 7$, but were rejected because the alternative low-redshift solution could not be ruled out at ≥95 per cent confidence. Further two candidates (ERS.YD7* and ERS.YD8*) were excluded because our primary photometric redshift solution lies at $z_{\rm phot} \simeq 2$. The remaining two objects (ERS.YD3 and ERS.YD4) do not appear as robust objects in any of our catalogues.

6.2.3 Wilkins et al. (2010)

Based on their analysis of the ERS field, Wilkins et al. (2010) identify a sample of 11 *z*-drop candidates, six of which also feature in our final robust sample. Of the five additional candidates listed by Wilkins et al., one object (ERS.z.26813) does have an acceptable primary photometric redshift solution at $z_{\rm phot} = 6.6$, but was excluded from our final sample because it has an equally acceptable solution at $z_{\rm phot} = 1.5$. A further object (ERS.z.70546) was rejected because it was not possible to obtain an acceptable high-redshift SED fit. The three remaining candidates listed by Wilkins et al. (ERS.z.80252, ERS.z.47667 and ERS.z.20851) were rejected as low-redshift interlopers by our SED-fitting analysis due to the presence of consistent, low-level, detections in the bluer optical bands. To illustrate this point, we have stacked the ACS+WFC3/IR data for these three objects and showed the resulting postage-stamp images in Fig. 9. It can clearly be seen that although there is a drop in flux between the z_{850} and Y_{098} filters, the significant detection of flux in the stack of the $B_{435} + V_{606} + i_{775}$ images suggests these objects are unlikely to be at $z \simeq 7$.

6.3 HUDF09-2

6.3.1 Bouwens et al. (2011)

Bouwens et al. (2011) lists a total of $N = 35$ robust high-redshift candidates in the HUDF09-2, consisting of 18 *z*-drops and 17 *Y*-drops. Only seven of these 35 candidates appear in our final robust sample (including HUDF09-2_799 which requires a contribution from Lyman α line emission), which clearly requires some explanation. The principal reason for this apparent discrepancy is that the Bouwens et al. (2011) sample contains many fainter objects than our final robust sample. Indeed, of the 35 candidates listed by Bouwens et al., 17 are fainter (in our 0.6-arcsec-diameter aperture photometry) than the faintest member of our final robust sample. Therefore, based on the data utilized in this study, and our criteria for isolating robust candidates, it is likely that these 17 objects are simply too faint to make it into our final robust sample.[8]

However, even accounting for the difference in selection depth, there are still 11 robust candidates identified by Bouwens et al. (2011) which should, in principle, also appear in our final robust sample. All 11 of these candidates do feature in our HUDF09-2 catalogues, but were excluded from the final robust sample for a number of reasons. Three of the additional candidates (UDF092z-00811320, UDF092z-07091160 and UDF092y-07090218) have acceptable high-redshift photometric redshift solutions, and were close to making it into our final robust candidate list. However, for these candidates, the difference in χ^2 between the primary photometric redshift

[7] This object was highlighted by Bouwens et al. (2011) as being potentially at low redshift.

[8] Bouwens et al. (2011) exploit deep F814W imaging which partially covers the HUDF09-2 field and, in some cases, will allow the selection of fainter high-redshift candidates.

solution and the alternative low-redshift solution ($\Delta\chi^2 \simeq 3$) did not quite match our adopted criterion of $\Delta\chi^2 \geq 4$. Of the remaining eight additional candidates listed by Bouwens et al. (2011), five (UDF092y-02731564, UDF092z-09770485, UDF092z-09151531, UDF092y-06321217 and UDF092y-06391247) were rejected because our primary photometric redshift solutions lie in the redshift interval $4.9 < z_{\rm phot} < 5.9$. The remaining three additional candidates (UDF092y-04242094, UDF092y-09611126 and UDF092y-09661163) were rejected because our analysis suggests that their primary photometric redshift solutions are at $z_{\rm phot} \simeq 2.1$.

Finally, we should note that two of the candidates which appear in our final robust sample (HUDF09-2_2455 and HUDF09-2_2814) also feature in the Bouwens et al. (2011) list of potential, but non-robust, high-redshift candidates. Moreover, our final robust sample features eight candidates which do not appear in any of the Bouwens et al. (2011) lists, although seven out of eight of these additional candidates have $z_{\rm phot} \leq 6.3$, where the z-drop criteria applied by Bouwens et al. are less sensitive.

6.3.2 *Wilkins et al. (2010)*

Wilkins et al. (2010) list a total of 15 z-drop candidates in the HUDF09-2, of which six also appear in our final robust sample. All nine of the additional candidates listed by Wilkins et al. (2010) feature in our initial HUDF09-2 catalogues, but were excluded from the final sample for a number of different reasons. One candidate (P34.z.3996) does have a valid high-redshift solution at $z_{\rm phot} = 6.55$, and was close to making it into the final robust sample, but was rejected because the $\Delta\chi^2$ between the high-redshift and low-redshift alternative solutions was too small ($\Delta\chi^2 = 3.4$). Of the remaining eight candidates, two candidates (P34.z.703 and P34.z.2428) were excluded because their primary photometric redshift solutions are in the interval $4.5 < z_{\rm phot} < 5.5$. The remaining six candidates (P34.z.2397, P34.z.3053, P34.z.4288, P34.z.4501, P34.z.5016 and P34.z.3990) were rejected because their primary photometric redshift solutions were all at $z_{\rm phot} \leq 4.5$. As an illustration, in Fig. 10, we show our SED fits for three of the z-drop candidates from Wilkins et al. (2010). In each case, our analysis suggests that there is a very low probability of the candidate being at $z \simeq 7$.

6.3.3 *Lorenzoni et al. (2011)*

Lorenzoni et al. (2011) list a total of seven Y-drop candidates in the HUDF09-2 (two highlighted with an * as more marginal candidates). None of these seven candidates makes it into our final robust sample. Four of the seven Lorenzoni et al. Y-drop candidates (P34.YD1, P34.YD2, P34.YD3 and P34.YD6*) feature in our original HUDF09-2 catalogues, and do have acceptable photometric redshift solutions at $7.0 < z_{\rm phot} < 8.5$. However, all four objects have comparable/preferred solutions in the redshift range $1.3 < z_{\rm phot} < 2.1$ and were therefore rejected. Of the remaining objects, one (P34.YD7*) is not a robust object in our reduction of the HUDF09-2 data set. The final two candidates were excluded because they were deemed to have unreliable photometry, either due to being within the wings of a bright star (P34.YD4) or due to being too close to the noisy array edge (P34.YD5).

7 CONCLUSIONS

We have presented the results of a study designed to identify robust high-redshift ($z \geq 6$) galaxies using the available multiwavelength

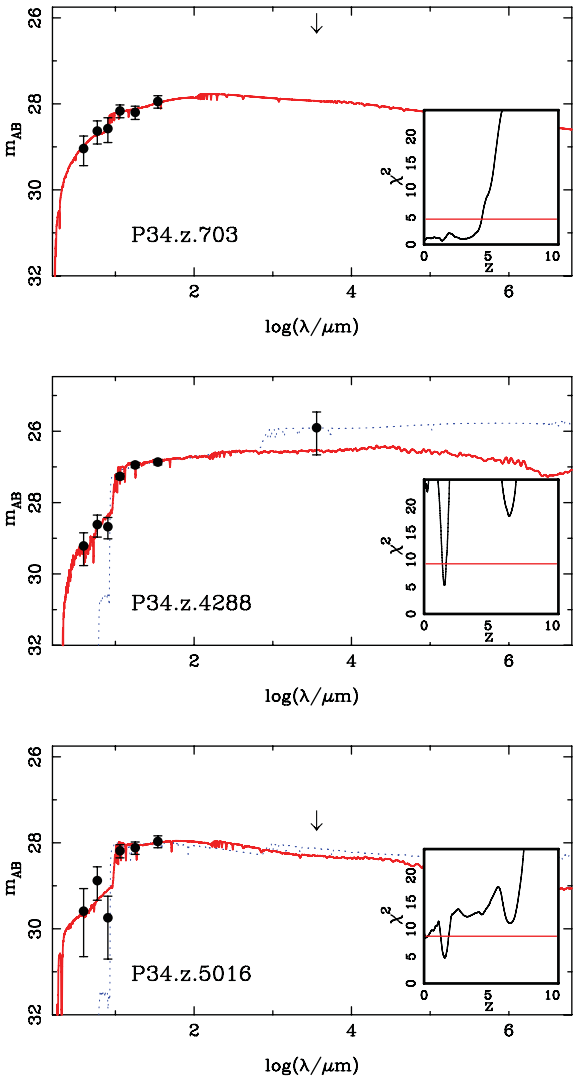

Figure 10. The results of our SED-fitting analysis for three $z \simeq 7$ candidates in the HUDF09-2 (P34) from Wilkins et al. (2010). Each of these objects was rejected from our final robust sample because our SED-fitting analysis does not return a plausible high-redshift solution. In each panel, the solid (red) line shows the best-fitting galaxy template at low redshift, while the dotted (blue) line shows the alternative high-redshift solution (note: there is no high-redshift solution for P34.z.703). The best-fitting photometric redshifts for these candidates are $z_{\rm phot} = 1.4$, 1.6 and 1.6 for P34.z.703, P34.z.4288 and P34.z.5016, respectively. In each panel, the inset showing χ^2 as a function of redshift demonstrates that any solution at $z \simeq 7$ has a very low probability.

(ACS+WFC3/IR+IRAC) imaging covering the HUDF, HUDF09-2 and ERS field. By exploiting the advantages provided by our SED-fitting analysis, and by incorporating deconfused IRAC photometry directly into our candidate-selection procedure, we have produced a robust sample of 70 high-redshift candidates at redshifts $6.0 < z_{\rm phot} < 8.7$. Based on this final robust sample, we have investigated the distribution of UV spectral slopes, stellar masses and SFRs. Finally, we have presented the results of a detailed comparison between our final robust sample and previous samples of high-redshift candidates derived from the same data sets using different selection criteria. Our main conclusions can be summarized as follows.

(i) Based on our final sample of 70 robust high-redshift candidates, and employing a variance-weighted mean to account for the

wide range in individual uncertainties, we find that the typical value of the UV spectral slope is $\langle\beta\rangle = -2.05 \pm 0.09$. Consequently, in contrast to some previous studies, we find no evidence that LBGs in the redshift interval $6.0 < z_{\rm phot} < 8.7$ display UV spectral slopes which are significantly bluer than seen in star-forming galaxies at lower redshifts.

(ii) Using the stellar-mass estimates provided by our SED fitting, we have investigated the relationship between stellar mass and UV luminosity at $z \simeq 7$, finding a best-fitting relationship of the form $M_\star \propto L_{1500}^{2.1\ \pm 0.6}$. We note that our best-fitting $M_\star$–$L_{\rm UV}$ relation is consistent with, although somewhat steeper than, the $M_\star$–$L_{\rm UV}$ relations derived by previous authors based on large samples of LBGs at $z \simeq 4$.

(iii) Focusing on a subsample of high-redshift candidates with reliable IRAC photometry, we find that $L^\star$ LBGs at $z \simeq 6.5$ (i.e. $M_{1500} \simeq -20.2$) have a median stellar mass of $M_\star = (2.1 \pm 1.1) \times 10^9\,{\rm M}_\odot$. However, by employing SED templates featuring a range of metallicities, SFHs and reddening, we find that the scatter in stellar mass at a given UV luminosity could span a factor of $\simeq$50. We caution that deriving stellar-mass estimates solely from models with a CSF may provide a misleadingly low estimate of the real range in stellar mass at a given UV luminosity.

(iv) Based on the same subsample of 21 objects with IRAC detections, we find that the median sSFR of $L^\star$ LBGs at $z \simeq 6.5$ is $1.9 \pm 0.8\ {\rm Gyr}^{-1}$, a value which is consistent with the sSFR $\simeq 2.5\ {\rm Gyr}^{-1}$ observed in star-forming galaxies at $z \simeq 2$. However, our SED fitting indicates that the range in the SFR at a given stellar mass is potentially large and that adopting CSF models with zero reddening may provide an underestimate of the real level of scatter. Moreover, we find that SED fitting based solely on CSF models (with zero reddening) will inevitably tend to produce a typical sSFR close to $\simeq 2.5\ {\rm Gyr}^{-1}$ for high-redshift galaxies with ages in the range 200–600 Myr.

(v) Using the subsample of 21 objects with IRAC detections, we have also investigated the impact on our stellar-mass estimates of including the effects of nebular continua and line emission in our galaxy SED templates. Based on SED templates with subsolar metallicity ($Z = 0.2\ {\rm Z}_\odot$), we find that nebular ($f_{\rm esc} = 0.2$) models are typically capable of providing a better fit to the predominantly blue $[3.6] - [4.5]\,\mu$m colours. However, in the majority of cases (16 out of 21), the best-fitting stellar masses returned by the nebular fits are less than a factor of 2.5 lower than those returned by the stellar continuum fits. Those objects for which the nebular fits return significantly lower stellar masses (5 out of 21) are found to have a somewhat unphysical median age of 9 Myr.

(vi) A detailed comparison between our final robust sample and previous high-redshift samples derived using different selection/analysis techniques produces mixed results. When confined to the brightest objects, and the best data, the overlap between our final robust sample and samples derived using traditional 'drop-out' criteria is reasonably good. However, at fainter magnitudes, our analysis suggests that some literature samples are heavily contaminated (in some cases at the $\geq$50 per cent level) by objects which are very likely at lower redshifts.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous referee whose comments and suggestions significantly improved the final version of this manuscript. RJM would like to acknowledge the funding of the Royal Society via the award of a University Research Fellowship, and the Leverhulme Trust via the award of a Philip Leverhulme research prize. JSD acknowledges the support of the Royal Society via a Wolfson Research Merit award, and also the support of the European Research Council via the award of an Advanced Grant. LdR acknowledges STFC funding via the award of a STFC PDRA. MC acknowledges the award of an STFC Advanced Fellowship. BER is supported by a Hubble Fellowship grant, programme number HST-HF-51262.01-A, provided by NASA from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. DPS acknowledges support from the STFC through the award of a Postdoctoral Research Fellowship. RAAB acknowledges the support provided by a European Research Council Advanced Grant PhD studentship. The authors would like to acknowledge Stéphane Charlot for providing the CB07 stellar population models.

REFERENCES

Anders P., Fritze-v. Alvensleben U., 2003, A&A, 401, 1063
Becker G. D., Rauch M., Sargent W. L. W., 2007, ApJ, 662, 72
Beckwith S. V. W. et al., 2006, AJ, 132, 1729
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bouwens R. J., Illingworth G. D., Franx M., Ford H., 2007, ApJ, 670, 928
Bouwens R. J. et al., 2010a, ApJ, 709, L133
Bouwens R. J. et al., 2010b, ApJ, 708, L69
Bouwens R. J. et al., 2011, ApJ, 737, 90
Brinchmann J. et al., 2004, MNRAS, 351, 1151
Brown R. L., Mathews W. G., 1970, ApJ, 160, 939
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000 (BC03)
Bunker A. J. et al., 2010, MNRAS, 409, 855
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Cardamone C. N. et al., 2010, ApJS, 189, 270
Chabrier G., 2003, PASP, 115, 763
Daddi E. et al., 2007, ApJ, 670, 156
Damen M. et al., 2011, ApJ, 727, 1
De Santis C., Grazian A., Fontana A., Santini P., 2007, New Astron., 12, 271
Dunkley J. et al., 2009, ApJ, 180, 306
Dunlop J. S., McLure R. J., Robertson B. E., Ellis R. S., Stark D. P., Cirasuolo M., de Ravel L., 2011, MNRAS, preprint (arXiv:1102.5005)
Elbaz D. et al., 2007, A&A, 468, 33
Fan X. et al., 2006, AJ, 132, 117
Finkelstein S. L., Papovich C., Giavalisco M., Reddy N. A., Ferguson H. C., Koekemoer A. M., Dickinson M., 2010, ApJ, 719, 1250
Finkelstein S. L. et al., 2011, ApJ, 729, 140
Finlator K., Oppenheimer B. D., Davé R., 2011, MNRAS, 410, 1703
Giavalisco M. et al., 2004, ApJ, 600, L93
González V., Labbé I., Bouwens R. J., Illingworth G., Franx M., Kriek M., Brammer G. B., 2010, ApJ, 713, 115
González V., Labbé I., Bouwens R. J., Illingworth G., Franx M., Kriek M., 2011, ApJ, 735, L34
Grazian A. et al., 2011, A&A, 532, 33
Grogin N. A. et al., 2011, ApJS, preprint (arXiv:1105.3753)
Högbom J., 1974, ApJS, 15, 417
Karim A. et al., 2011, ApJ, 730, 61
Koekemoer A. M., Fruchter A. S., Hook R. N., Hack W. I., 2002, in Arribas S., Koekemoer A. M., Whitmore B., eds, The 2002 *HST* Calibration Workshop. Space Telescope Science Institute, Baltimore, p. 337
Koekemoer A. M. et al., 2011, ApJS, preprint (arXiv:1105.3754)
Labbé I., Bouwens R., Illingworth G. D., Franx M., 2006, ApJ, 649, L67
Labbé I. et al., 2010, ApJ, 716, L103
Laidler V. G. et al., 2007, PASP, 119, 1325
Lehnert M. D. et al., 2010, Nat, 467, L940

Lorenzoni S., Bunker A., Wilkins S., Stanway E., Jarvis M., Caruana J., 2011, MNRAS, 414, 1455
McLure R. J. et al., 2006, MNRAS, 372, 357
McLure R. J., Cirasuolo M., Dunlop J. S., Foucaud S., Almaini O., 2009, MNRAS, 395, 2196
McLure R. J., Dunlop J. S., Cirasuolo M., Koekemoer A. M., Sabbi E., Stark D. P., Target T. A., Ellis R. S., 2010, MNRAS, 403, 960
Madau P., 1995, ApJ, 441, 18
Madau P., Pozzetti L., Dickinson M., 1998, ApJ, 498, 106
Magdis G. E., Rigopoulou D., Huang J. S., Fazio G. G., 2010, MNRAS, 401, 1521
Nakamura E., Inoue A. K., Hayashino T., Horie M., Kousai K., Fujii T., Matsuda Y., 2011, MNRAS, 412, 2579
Noeske K. G. et al., 2007, ApJ, 660, L43
Oesch P. et al., 2007, ApJ, 671, 1212
Oesch P. et al., 2010, ApJ, 709, L16
Oke J. B., Gunn J. E., 1983, ApJ, 266, 713
Ono Y., Ouchi M., Shimasaku K., Dunlop J., Farrah D., McLure R., Okamura S., 2010, ApJ, 724, 1524
Osterbrock D. E., Ferland G. J., 2006, Astrophysics of Gaseous Nebulae and Active Galactic Nuclei, 2nd edn. University Science Books, Mill Valley, CA
Ota K. et al., 2010, ApJ, 722, 803
Ouchi M. et al., 2009, ApJ, 706, 1136
Ouchi M. et al., 2010, ApJ, 723, 869
Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical Recipes. Cambridge Univ. Press, Cambridge
Robertson B. E., Ellis R. S., Dunlop J. S., McLure R. J., Stark D. P., 2010, Nat, 468, 49
Salpeter E. E., 1955, ApJ, 121, 161
Schaerer D., de Barros S., 2010, A&A, 515, A73
Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., Erb D. K., 2006, ApJ, 651, 688
Stark D. P., Ellis R. S., Richard S., Bunker A., Bundy K., Targett T., Benson A., Lacy M., 2009, ApJ, 697, 1493
Stark D. P., Ellis R. S., Kuenley C., Ouchi M., Bunker A., 2010, MNRAS, 408, 1628
Szalay A. S., Connolly A. J., Szokoly G. P., 1999, AJ, 117, 68
Vanzella E. et al., 2011, ApJ, 730, L35
Wang W. H., Cowie L. L., Barger A. J., Keenan R. C., Ting H. C., 2010, ApJS, 187, 251
Wilkins S. M., Bunker A. J., Lorenzoni S., Caruana J., 2010, MNRAS, preprint (arXiv:1002.4866)
Windhorst R. A. et al., 2011, ApJS, 193, 27
Yan H. J., Windhorst R. A., Hathi N. P., Cohen S. H., Ryan R. E., O'Connell R. W., McCarthy P. J., 2010, Res. Astron. Astrophys., 10, 867

APPENDIX A: IRAC DECONFUSION ALGORITHM

As previously discussed in Section 2.3, deep *Spitzer* IRAC imaging data are available for all three of the fields analysed in this paper (3.6 + 4.5 μm for the HUDF+ERS field and 3.6 μm for the HUDF09-2) with total integration times ranging from 23 to 46 h. Fully exploiting the information provided by the IRAC imaging is vital for analysing the high-redshift galaxy population for two fundamental reasons. First, given that even the H_{160} imaging data are sampling rest-frame wavelengths of $\lambda \leq 2200$ Å at $z \geq 6$, the information longwards of the 4000-Å break provided by the IRAC imaging is crucial for providing constraints on the stellar mass and the contribution of any older stellar population (see Section 5). Secondly, inclusion of the IRAC photometry in the SED-fitting process is very effective at breaking photometric redshift degeneracies between genuine $z \geq 6$ candidates and interlopers at both $z \simeq 2$ and $z \simeq 5$ (see SED plots in Appendix B).

Although the availability of ultradeep IRAC imaging is potentially hugely beneficial, because the IRAC data are heavily confused, obtaining accurate flux measurements for faint, high-redshift galaxies is technically challenging. Several approaches to solving this problem have been presented in the literature, all of which rely on using a higher resolution image (ideally as close as possible in wavelength) as prior information to deconfuse the lower resolution IRAC data. One approach, which has been recently applied to the ultradeep IRAC data in the GOODS-N field, uses a modified version of the CLEAN algorithm (Högbom 1974), traditionally used in radio astronomy, together with model templates extracted from high-resolution, ground-based, *K*-band imaging to estimate IRAC fluxes via an iterative subtraction scheme (Wang et al. 2010). An alternative approach relies on building two-dimensional, axisymmetric models of each galaxy based on the available high-resolution imaging (e.g. Labbé et al. 2006). The two-dimensional models are then convolved to the spatial resolution of the IRAC imaging and a χ^2-minimization procedure is employed to determine the individual galaxy fluxes which best reproduce the observed IRAC image. A distinct advantage of this approach is that it allows the precise centroiding of the individual galaxies to be included as a free parameter in the fitting process, which can compensate for slight astrometric differences between the high-resolution and IRAC images. However, this method obviously has the disadvantage of relying on axisymmetric galaxy models, and is only really suitable for deconfusing relatively small areas of IRAC imaging at a time.

The method adopted in this paper relies on a related, but different, approach, whereby the actual two-dimensional light distributions of individual galaxies in the *HST* imaging are used as the model templates. In this approach, SExtractor is used to produce a normalized template of each individual galaxy based on either J_{125} or H_{160} WFC3/IR imaging, and is transformed into IRAC spatial resolution via convolution with a transfer function:

$$W_{\rm PSF} * T = I_{\rm PSF}, \tag{A1}$$

where T is the transfer function, and $W_{\rm PSF}$ and $I_{\rm PSF}$ are the WFC3/IR and IRAC PSFs, respectively. The fundamental assumption behind this technique is that it is possible to reproduce the observed IRAC image using a linear combination of these galaxy templates. The key advantage of this approach is that, provided the astrometry match between the WFC3/IR and IRAC images is sufficiently accurate that the centroid of each galaxy template can be held fixed, the amplitude of each template can be uniquely (and analytically) determined by χ^2 minimization as follows:

$$\chi^2 = \sum_{ij} \frac{\left(D_{ij} - \sum_{k=1}^{n} a_k M_{ij}^k\right)^2}{\sigma_{ij}^2}, \tag{A2}$$

where D_{ij} is the (i, j)th pixel of the IRAC image, σ_{ij} is the corresponding uncertainty and M_{ij}^k is the (i, j)th pixel of the kth galaxy template. The minimum χ^2 occurs when

$$\frac{\delta\chi^2}{\delta a_k} = 2\sum_{ij} \frac{\left(D_{ij} - \sum_{k=1}^{n} a_k M_{ij}^k\right) M_{ij}^m}{\sigma_{ij}^2} = 0, \tag{A3}$$

which can be rearranged as

$$\sum_{k=1}^{n} \left(\sum_{ij} \frac{M_{ij}^k M_{ij}^m}{\sigma_{ij}^2} \right) a_k = \sum_{ij} \frac{D_{ij} M_{ij}^m}{\sigma_{ij}^2}, \tag{A4}$$

which describes a set of n linear equations, which can be recast as a matrix equation

$$\mathbf{A}\boldsymbol{a} = \boldsymbol{b}, \tag{A5}$$

Table B1. Basic observational properties of the final high-redshift galaxy sample in the HUDF. Columns one to three list the candidate IDs and coordinates. The remaining columns list the photometry of each candidate in the i_{775}, z_{850}, Y_{105}, H_{160}, IRAC1 (3.6-μm) and IRAC2 (4.5-μm) filters, as measured in a 0.6-arcsec-diameter aperture, along with their corresponding uncertainties. The magnitudes listed here are derived from the actual fluxes used in the SED fitting and are not corrected to total, but have been corrected for galactic extinction and the relative aperture losses between the ACS and WFC3/IR. All candidates are undetected in filters at shorter wavelengths than i_{775}, and all detections which are significant at less than the 2σ level are listed as 2σ upper limits. It should be noted that because the IRAC photometry is derived via a deconfusion process, the uncertainties and 2σ limits are highly position-dependent.

ID	RA (J2000)	Dec. (J2000)	i_{775}	z_{850}	Y_{105}	J_{125}	H_{160}	IRAC1	IRAC2
HUDF_1344	03:32:36.63	−27:47:50.1	>30.28	$27.85^{+0.16}_{-0.14}$	$27.66^{+0.11}_{-0.10}$	$27.70^{+0.09}_{-0.08}$	$27.77^{+0.09}_{-0.08}$	>26.83	>26.78
HUDF_1016	03:32:35.06	−27:47:40.2	>30.21	$27.83^{+0.10}_{-0.09}$	$27.49^{+0.08}_{-0.07}$	$27.20^{+0.08}_{-0.07}$	$27.27^{+0.08}_{-0.07}$	$26.94^{+0.37}_{-0.28}$	>27.38
HUDF_522	03:32:36.47	−27:46:41.4	$28.88^{+0.12}_{-0.11}$	$26.54^{+0.08}_{-0.07}$	$26.18^{+0.08}_{-0.07}$	$26.16^{+0.08}_{-0.07}$	$26.20^{+0.08}_{-0.07}$	$25.62^{+0.24}_{-0.20}$	$26.45^{+0.45}_{-0.32}$
HUDF_2622	03:32:36.64	−27:47:50.2	>30.28	$28.45^{+0.30}_{-0.23}$	$28.21^{+0.18}_{-0.15}$	$28.24^{+0.14}_{-0.13}$	$28.33^{+0.14}_{-0.13}$	>26.64	>26.52
HUDF_796	03:32:37.46	−27:46:32.8	>29.90	$27.72^{+0.08}_{-0.07}$	$27.08^{+0.08}_{-0.07}$	$27.02^{+0.08}_{-0.07}$	$26.88^{+0.08}_{-0.07}$	>25.36	$25.00^{+0.47}_{-0.33}$
HUDF_2836	03:32:35.05	−27:47:25.8	>30.55	$29.03^{+0.38}_{-0.28}$	$28.44^{+0.25}_{-0.20}$	$28.92^{+0.43}_{-0.31}$	$28.99^{+0.24}_{-0.19}$	>26.64	>26.52
HUDF_1692	03:32:43.03	−27:46:23.6	>30.07	$28.03^{+0.17}_{-0.15}$	$27.62^{+0.11}_{-0.10}$	$27.88^{+0.11}_{-0.10}$	$28.14^{+0.16}_{-0.14}$	>27.26	>27.03
HUDF_2743	03:32:36.52	−27:46:42.0	>30.21	$29.37^{+0.58}_{-0.38}$	$28.90^{+0.33}_{-0.25}$	$28.75^{+0.28}_{-0.22}$	$29.01^{+0.42}_{-0.30}$	>26.64	>26.52
HUDF_2316	03:32:44.31	−27:46:45.2	>30.76	$28.89^{+0.38}_{-0.28}$	$28.38^{+0.28}_{-0.22}$	$28.33^{+0.20}_{-0.17}$	$28.74^{+0.34}_{-0.26}$	>27.14	>26.97
HUDF_2281	03:32:39.79	−27:46:33.7	>29.90	$29.10^{+0.37}_{-0.28}$	$28.40^{+0.16}_{-0.14}$	$28.62^{+0.30}_{-0.23}$	$28.70^{+0.25}_{-0.21}$	>27.04	>26.86
HUDF_1442	03:32:42.19	−27:46:27.8	>30.01	$28.77^{+0.26}_{-0.21}$	$27.93^{+0.10}_{-0.09}$	$27.83^{+0.09}_{-0.09}$	$27.87^{+0.08}_{-0.07}$	>26.83	$26.13^{+0.39}_{-0.29}$
HUDF_2324	03:32:41.60	−27:47:04.5	>29.95	$29.18^{+0.35}_{-0.26}$	$28.39^{+0.18}_{-0.16}$	$28.55^{+0.20}_{-0.17}$	$28.75^{+0.23}_{-0.19}$	>27.07	>26.92
HUDF_2672	03:32:37.80	−27:47:40.4	>30.21	$29.18^{+0.49}_{-0.34}$	$28.38^{+0.20}_{-0.17}$	$28.60^{+0.23}_{-0.19}$	$28.59^{+0.21}_{-0.17}$	>26.64	>26.52
HUDF_1818	03:32:36.38	−27:47:16.3	>30.21	$29.18^{+0.47}_{-0.33}$	$27.95^{+0.12}_{-0.11}$	$28.22^{+0.20}_{-0.17}$	$28.28^{+0.14}_{-0.13}$	>27.19	>27.02
HUDF_1473	03:32:36.77	−27:47:53.6	>29.90	$28.98^{+0.41}_{-0.29}$	$27.78^{+0.09}_{-0.09}$	$27.96^{+0.10}_{-0.09}$	$27.91^{+0.08}_{-0.07}$	>27.34	>27.14
HUDF_1730	03:32:43.78	−27:46:33.7	>30.07	$29.16^{+0.48}_{-0.33}$	$28.04^{+0.12}_{-0.11}$	$28.03^{+0.10}_{-0.09}$	$28.18^{+0.17}_{-0.15}$	>26.78	>26.65
HUDF_1632	03:32:37.44	−27:46:51.2	>30.07	$29.10^{+0.32}_{-0.24}$	$27.97^{+0.10}_{-0.09}$	$28.06^{+0.12}_{-0.11}$	$28.08^{+0.14}_{-0.13}$	>27.21	>27.04
HUDF_2084	03:32:40.57	−27:46:43.6	>30.45	$29.57^{+0.50}_{-0.34}$	$28.19^{+0.18}_{-0.16}$	$28.49^{+0.16}_{-0.14}$	$28.54^{+0.14}_{-0.13}$	>27.42	>27.17
HUDF_1995	03:32:39.58	−27:46:56.5	>30.01	$29.29^{+0.72}_{-0.43}$	$28.24^{+0.20}_{-0.17}$	$28.12^{+0.18}_{-0.15}$	$28.45^{+0.15}_{-0.13}$	>27.00	>26.90
HUDF_658	03:32:42.56	−27:46:56.6	>30.37	$28.13^{+0.20}_{-0.17}$	$26.73^{+0.08}_{-0.07}$	$26.53^{+0.08}_{-0.07}$	$26.54^{+0.08}_{-0.07}$	$25.65^{+0.24}_{-0.20}$	$26.06^{+0.29}_{-0.23}$
HUDF_2701	03:32:41.82	−27:46:11.3	>30.14	$29.79^{+0.70}_{-0.42}$	$28.62^{+0.31}_{-0.24}$	$28.66^{+0.23}_{-0.19}$	$28.96^{+0.30}_{-0.24}$	>26.64	>26.52
HUDF_860	03:32:38.81	−27:47:07.2	>30.55	>29.75	$27.61^{+0.10}_{-0.09}$	$27.14^{+0.08}_{-0.07}$	$27.01^{+0.08}_{-0.07}$	$25.69^{+0.24}_{-0.20}$	$26.12^{+0.26}_{-0.21}$
HUDF_1102	03:32:39.55	−27:47:17.5	>29.84	>29.49	$28.08^{+0.15}_{-0.13}$	$27.53^{+0.08}_{-0.07}$	$27.42^{+0.09}_{-0.09}$	>26.30	>26.20
HUDF_1419	03:32:43.13	−27:46:28.5	>30.28	>29.75	$28.34^{+0.16}_{-0.14}$	$28.00^{+0.13}_{-0.12}$	$27.84^{+0.09}_{-0.09}$	>26.69	>26.62
HUDF_2641	03:32:39.73	−27:46:21.3	>30.14	>29.39	$28.96^{+0.38}_{-0.28}$	$28.41^{+0.19}_{-0.16}$	$28.92^{+0.29}_{-0.23}$	>26.64	>26.52
HUDF_1962	03:32:38.36	−27:46:11.9	>30.21	>29.53	$28.69^{+0.33}_{-0.25}$	$28.14^{+0.20}_{-0.17}$	$28.42^{+0.17}_{-0.15}$	>26.55	>26.61
HUDF_1173	03:32:44.70	−27:46:44.3	$30.45^{+0.75}_{-0.44}$	>29.57	$28.07^{+0.31}_{-0.24}$	$27.45^{+0.10}_{-0.10}$	$27.53^{+0.10}_{-0.09}$	$26.92^{+0.66}_{-0.41}$	>26.87
HUDF_2664	03:32:33.13	−27:46:54.5	>29.95	>29.53	$29.18^{+0.43}_{-0.31}$	$28.55^{+0.26}_{-0.21}$	$29.18^{+0.63}_{-0.40}$	>26.64	>26.52
HUDF_1660	03:32:37.21	−27:48:06.2	>30.14	>29.75	$28.59^{+0.20}_{-0.17}$	$27.99^{+0.09}_{-0.09}$	$28.12^{+0.13}_{-0.12}$	>27.24	>27.01
HUDF_1679	03:32:42.88	−27:46:34.5	>30.14	>29.49	$29.14^{+0.24}_{-0.20}$	$28.14^{+0.16}_{-0.14}$	$28.13^{+0.10}_{-0.09}$	>27.32	>27.04
HUDF_2003	03:32:38.13	−27:45:54.0	>30.28	>29.79	>29.66	$28.40^{+0.20}_{-0.17}$	$28.46^{+0.19}_{-0.16}$	>27.38	>27.10

where

$$\mathbf{A}_{mk} = \sum_{ij} \frac{M_{ij}^m M_{ij}^k}{\sigma_{ij}^2} \quad \text{and} \quad b_k = \sum_{ij} \frac{D_{ij} M_{ij}^k}{\sigma_{ij}^2}.$$

From equation (A5), it is possible to find $\boldsymbol{a}$, the n-dimensional vector of galaxy template amplitudes, using standard matrix inversion techniques. Moreover, the variance of the fitted amplitudes is automatically provided by the diagonal terms of the inverse matrix:

$$\sigma^2(a_k) = \mathbf{A}_{kk}^{-1}. \tag{A6}$$

We should note that the basic algorithm outlined here is very similar to that employed by the publicly available TFIT (Laidler et al. 2007) and CONVPHOT (De Santis et al. 2007) software packages.

APPENDIX B: CANDIDATE PHOTOMETRY, SED FITS AND POSTAGE-STAMP IMAGES

In Tables B1–B3, we provide the photometry for each member of our final robust sample. In Figs B1–B3, we illustrate the results of our SED-fitting analysis by showing the best-fitting high-redshift galaxy template, the best-fitting low-redshift alternative solution and the dependence of χ^2 on photometric redshift. This information is included to allow the readers to judge for themselves the robustness of each high-redshift candidate. In Fig. B4, we show the results of fitting the 21-object subsample detected at 3.6 μm with SED models featuring nebular continua and line emission. Finally, in Figs B5–B7, we also provide 3×3 arcsec^2 postage-stamp images of each high-redshift candidate in the z_{850}, Y_{105}/Y_{098}, J_{125} and H_{160} filters.

Table B2. Basic observational properties of the final high-redshift galaxy sample in the ERS field. Columns one to three list the candidate IDs and coordinates. The remaining columns list the photometry of each candidate in the i_{775}, z_{850}, Y_{098}, H_{160}, IRAC1 (3.6-μm) and IRAC2 (4.5-μm) filters, as measured in a 0.6-arcsec-diameter aperture, along with their corresponding uncertainties. The magnitudes listed here are derived from the actual fluxes used in the SED fitting and are not corrected to total, but have been corrected for galactic extinction and the relative aperture losses between the ACS and WFC3/IR. All candidates are undetected in filters at shorter wavelengths than i_{775}, and all detections which are significant at less than the 2σ level are listed as 2σ upper limits. It should be noted that because the IRAC photometry is derived via a deconfusion process, the uncertainties and 2σ limits are highly position-dependent.

ID	RA (J2000)	Dec. (J2000)	i_{775}	z_{850}	Y_{098}	J_{125}	H_{160}	IRAC1	IRAC2
ERS_7086	03:32:34.75	−27:40:35.1	>28.28	$27.25^{+0.21}_{-0.18}$	$26.78^{+0.12}_{-0.11}$	$26.89^{+0.12}_{-0.11}$	$26.99^{+0.17}_{-0.14}$	$26.82^{+0.63}_{-0.40}$	>26.79
ERS_6066	03:32:07.86	−27:42:17.8	>28.19	$27.27^{+0.48}_{-0.33}$	$26.55^{+0.15}_{-0.13}$	$26.68^{+0.08}_{-0.08}$	$26.50^{+0.13}_{-0.11}$	$24.86^{+0.24}_{-0.20}$	$24.99^{+0.24}_{-0.20}$
ERS_9869	03:32:15.40	−27:43:28.6	>28.04	$27.52^{+0.32}_{-0.25}$	$26.99^{+0.23}_{-0.19}$	$27.35^{+0.16}_{-0.14}$	$27.53^{+0.34}_{-0.26}$	>26.45	>26.39
ERS_8668	03:32:27.96	−27:41:19.0	>28.41	$27.63^{+0.48}_{-0.33}$	$27.11^{+0.15}_{-0.14}$	$27.17^{+0.16}_{-0.14}$	$27.56^{+0.36}_{-0.27}$	>26.65	>26.58
ERS_9100	03:32:20.24	−27:43:34.3	>28.45	$27.87^{+0.58}_{-0.38}$	$27.16^{+0.10}_{-0.09}$	$27.24^{+0.12}_{-0.11}$	$27.20^{+0.14}_{-0.13}$	$26.39^{+0.31}_{-0.24}$	$26.52^{+0.43}_{-0.31}$
ERS_7225	03:32:36.31	−27:40:15.0	>28.12	>27.88	$27.00^{+0.18}_{-0.15}$	$26.92^{+0.08}_{-0.08}$	$26.82^{+0.16}_{-0.14}$	$24.88^{+0.24}_{-0.20}$	$25.23^{+0.24}_{-0.20}$
ERS_6438	03:32:25:28	−27:43:24.2	>27.99	>27.82	$26.77^{+0.15}_{-0.13}$	$26.76^{+0.10}_{-0.09}$	$26.57^{+0.11}_{-0.10}$	$24.83^{+0.30}_{-0.23}$	$25.73^{+0.58}_{-0.38}$
ERS_6263	03:32:06.83	−27:44:22.2	>28.17	$27.42^{+0.40}_{-0.29}$	$26.71^{+0.17}_{-0.15}$	$26.73^{+0.14}_{-0.12}$	$26.95^{+0.18}_{-0.15}$	$26.25^{+0.39}_{-0.29}$	>26.77
ERS_7776	03:32:03.77	−27:44:54.4	>28.29	$27.95^{+0.50}_{-0.34}$	$27.12^{+0.20}_{-0.17}$	$27.03^{+0.12}_{-0.11}$	$27.28^{+0.16}_{-0.14}$	$26.36^{+0.31}_{-0.24}$	>26.98
ERS_5847	03:32:16.00	−27:43:01.4	>28.52	$27.68^{+0.28}_{-0.22}$	$26.60^{+0.10}_{-0.09}$	$26.63^{+0.08}_{-0.07}$	$26.73^{+0.19}_{-0.16}$	$25.61^{+0.69}_{-0.42}$	>25.65
ERS_8987	03:32:16.01	−27:41:59.0	>28.06	>28.16	$27.62^{+0.46}_{-0.32}$	$27.22^{+0.16}_{-0.14}$	$27.17^{+0.19}_{-0.16}$	>25.36	$24.58^{+0.35}_{-0.26}$
ERS_3679	03:32:22.66	−27:43:00.7	>28.24	$27.28^{+0.22}_{-0.18}$	$26.19^{+0.09}_{-0.08}$	$25.94^{+0.08}_{-0.07}$	$25.91^{+0.08}_{-0.07}$	$25.07^{+0.26}_{-0.21}$	$25.67^{+0.50}_{-0.34}$
ERS_7412	03:32:09.85	−27:43:24.0	>28.18	>27.87	$26.76^{+0.13}_{-0.11}$	$26.96^{+0.10}_{-0.10}$	$26.66^{+0.12}_{-0.11}$	$26.89^{+0.75}_{-0.44}$	>26.61
ERS_6427	03:32:24.09	−27:42:13.9	>28.38	>27.89	$27.14^{+0.26}_{-0.21}$	$26.76^{+0.10}_{-0.09}$	$26.87^{+0.09}_{-0.08}$	$25.67^{+0.30}_{-0.23}$	$25.92^{+0.26}_{-0.21}$
ERS_8858	03:32:16.19	−27:41:49.8	>28.12	>27.70	$27.47^{+0.36}_{-0.27}$	$27.20^{+0.14}_{-0.13}$	$27.40^{+0.14}_{-0.13}$	>27.15	>26.94
ERS_7376	03:32:29.54	−27:42:04.5	>28.07	>28.19	$27.17^{+0.23}_{-0.19}$	$26.95^{+0.10}_{-0.09}$	$27.05^{+0.27}_{-0.22}$	$26.44^{+0.69}_{-0.42}$	>26.41
ERS_8176	03:32:23.15	−27:42:04.7	>28.01	>28.25	$27.06^{+0.22}_{-0.19}$	$27.09^{+0.13}_{-0.12}$	$27.50^{+0.24}_{-0.19}$	>27.06	>26.83
ERS_7672	03:32:10.03	−27:45:24.6	>28.06	>28.01	$27.07^{+0.23}_{-0.19}$	$27.01^{+0.21}_{-0.18}$	$27.10^{+0.14}_{-0.12}$	>26.94	>26.60
ERS_7475	03:32:32.81	−27:42:38.5	>28.49	>28.28	$27.42^{+0.37}_{-0.28}$	$26.97^{+0.11}_{-0.10}$	$26.94^{+0.11}_{-0.10}$	>26.47	>26.34
ERS_7236	03:32:11.51	−27:45:17.1	>28.07	>27.96	$27.67^{+0.34}_{-0.26}$	$26.93^{+0.14}_{-0.12}$	$27.25^{+0.13}_{-0.12}$	>26.40	>26.19
ERS_9041	03:32:23.37	−27:43:26.5	>28.39	>27.84	>28.48	$27.23^{+0.10}_{-0.09}$	$27.91^{+0.38}_{-0.28}$	>27.01	>26.82
ERS_10288	03:32:35.44	−27:41:32.7	>28.46	>27.97	>28.76	$27.40^{+0.12}_{-0.11}$	$27.47^{+0.16}_{-0.14}$	>27.19	>26.89
ERS_8584	03:32:02.99	−27:43:51.9	>28.16	>28.41	>28.52	$27.16^{+0.22}_{-0.18}$	$27.08^{+0.17}_{-0.14}$	>26.32	$25.79^{+0.54}_{-0.36}$
ERS_8496	03:32:29.69	−27:40:49.9	>28.63	>27.81	$27.18^{+0.27}_{-0.21}$	$27.14^{+0.18}_{-0.15}$	$27.41^{+0.33}_{-0.25}$	$25.36^{+0.24}_{-0.20}$	$25.91^{+0.37}_{-0.28}$
ERS_9923	03:32:10.06	−27:45:22.6	>28.00	>27.92	$26.82^{+0.20}_{-0.17}$	$27.36^{+0.28}_{-0.22}$	$27.29^{+0.16}_{-0.14}$	>26.95	>26.71

Table B3. Basic observational properties of the final high-redshift galaxy sample in the HUDF09-2. Columns one to three list the candidate IDs and coordinates. The remaining columns list the photometry of each candidate in the i_{775}, z_{850}, Y_{105}, H_{160} and IRAC1 (3.6 μm) filters, as measured in a 0.6-arcsec-diameter aperture, along with their corresponding uncertainties. The magnitudes listed here are derived from the actual fluxes used in the SED fitting and are not corrected to total, but have been corrected for galactic extinction and the relative aperture losses between the ACS and WFC3/IR. All candidates are undetected in filters at shorter wavelengths than i_{775}, and all detections which are significant at less than the 2σ level are listed as 2σ upper limits. It should be noted that because the IRAC photometry is derived via a deconfusion process, the uncertainties and 2σ limits are highly position-dependent.

ID	RA (J2000)	Dec. (J2000)	i_{775}	z_{850}	Y_{105}	J_{125}	H_{160}	IRAC1
HUDF09-2_2459	03:33:06.30	−27:50:20.2	>30.07	$27.88^{+0.15}_{-0.13}$	$27.83^{+0.16}_{-0.14}$	$27.73^{+0.10}_{-0.09}$	$27.94^{+0.12}_{-0.11}$	>25.08
HUDF09-2_2613	03:33:06.52	−27:50:34.6	>29.14	$27.92^{+0.20}_{-0.17}$	$27.74^{+0.13}_{-0.11}$	$27.91^{+0.15}_{-0.13}$	$28.20^{+0.21}_{-0.18}$	>25.89
HUDF09-2_2638	03:33:06.65	−27:50:30.2	>29.17	$28.50^{+0.33}_{-0.25}$	$28.02^{+0.16}_{-0.14}$	$28.03^{+0.14}_{-0.12}$	$28.06^{+0.16}_{-0.14}$	>26.17
HUDF09-2_1543	03:33:01.18	−27:51:22.3	>29.02	$27.08^{+0.08}_{-0.07}$	$26.62^{+0.08}_{-0.07}$	$26.66^{+0.08}_{-0.07}$	$26.77^{+0.08}_{-0.07}$	$26.31^{+0.42}_{-0.30}$
HUDF09-2_605	03:33:01.95	−27:52:03.2	>29.07	$28.20^{+0.23}_{-0.19}$	$27.53^{+0.10}_{-0.09}$	$27.51^{+0.08}_{-0.08}$	$27.56^{+0.11}_{-0.10}$	>24.87
HUDF09-2_2587	03:33:04.20	−27:50:31.3	>29.84	$27.47^{+0.09}_{-0.09}$	$26.73^{+0.08}_{-0.07}$	$26.60^{+0.08}_{-0.07}$	$26.63^{+0.08}_{-0.07}$	$25.71^{+0.45}_{-0.32}$
HUDF09-2_1660	03:33:01.10	−27:51:16.0	$29.12^{+0.73}_{-0.43}$	$27.65^{+0.16}_{-0.14}$	$26.82^{+0.08}_{-0.07}$	$26.64^{+0.08}_{-0.07}$	$26.71^{+0.08}_{-0.07}$	$25.71^{+0.41}_{-0.30}$
HUDF09-2_1745	03:33:01.19	−27:51:13.3	>29.14	$28.90^{+0.57}_{-0.37}$	$27.70^{+0.13}_{-0.12}$	$27.74^{+0.09}_{-0.09}$	$27.72^{+0.11}_{-0.10}$	>25.80
HUDF09-2_1620	03:33:05.40	−27:51:18.8	>29.07	>28.78	$27.88^{+0.14}_{-0.13}$	$28.15^{+0.22}_{-0.19}$	$28.07^{+0.13}_{-0.12}$	>24.37
HUDF09-2_1721	03:33:01.17	−27:51:13.9	>29.20	$28.75^{+0.46}_{-0.32}$	$27.69^{+0.15}_{-0.13}$	$27.23^{+0.08}_{-0.07}$	$27.24^{+0.08}_{-0.07}$	>25.26
HUDF09-2_2455	03:33:09.65	−27:50:50.8	>29.20	$28.17^{+0.16}_{-0.14}$	$26.58^{+0.08}_{-0.07}$	$26.64^{+0.08}_{-0.07}$	$26.62^{+0.08}_{-0.07}$	>22.60
HUDF09-2_1584	03:33:03.79	−27:51:20.4	>29.29	>28.80	$27.20^{+0.10}_{-0.09}$	$26.67^{+0.08}_{-0.07}$	$26.58^{+0.08}_{-0.07}$	$24.96^{+0.44}_{-0.31}$
HUDF09-2_2814	03:33:07.05	−27:50:55.5	>28.88	>28.86	$28.09^{+0.24}_{-0.20}$	$27.56^{+0.09}_{-0.08}$	$27.67^{+0.09}_{-0.08}$	>26.21
HUDF09-2_1596	03:33:03.76	−27:51:19.7	>29.39	>28.95	$27.35^{+0.09}_{-0.08}$	$26.83^{+0.08}_{-0.07}$	$26.82^{+0.08}_{-0.07}$	>24.52
HUDF09-2_2000	03:33:04.64	−27:50:53.0	>29.70	>29.79	$28.24^{+0.26}_{-0.21}$	$27.53^{+0.08}_{-0.07}$	$27.58^{+0.10}_{-0.09}$	>26.75
HUDF09-2_2765	03:33:07.58	−27:50:55.0	>29.29	>28.97	>29.26	$27.85^{+0.12}_{-0.11}$	$27.52^{+0.08}_{-0.07}$	>27.18
HUDF09-2_799	03:33:09.15	−27:51:55.4	>29.23	>29.06	$27.45^{+0.09}_{-0.08}$	$27.72^{+0.14}_{-0.13}$	$27.63^{+0.16}_{-0.13}$	>25.69

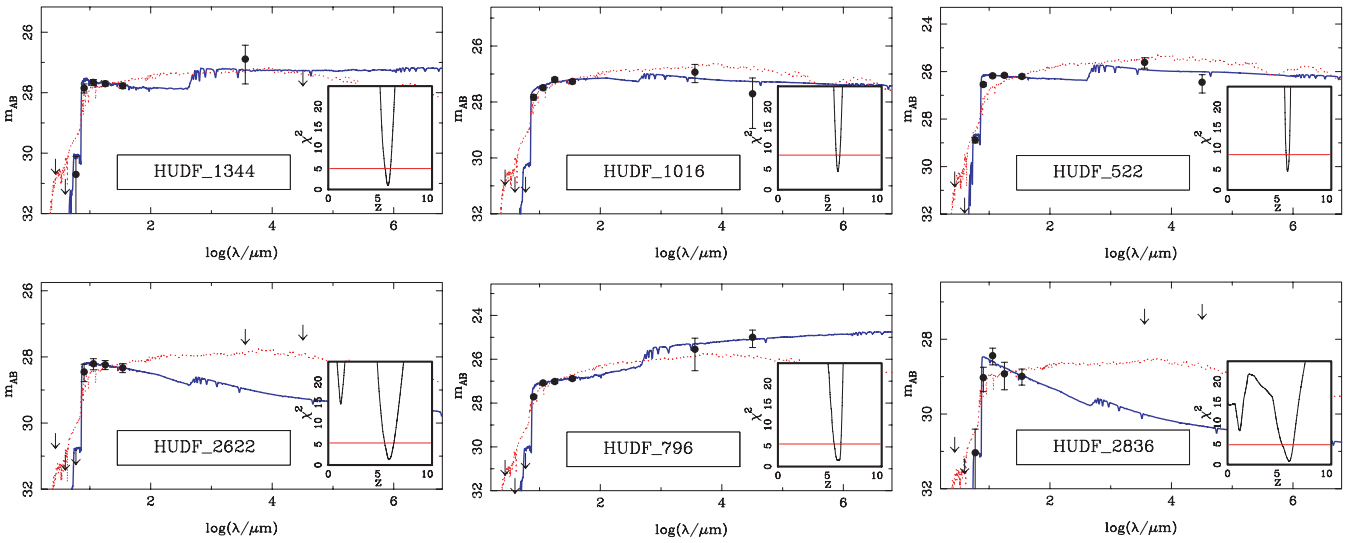

Figure B1. SED fits for each member of the final HUDF subsample. In each plot, the solid (blue) line is the best-fitting $z \geq 6$ galaxy SED template, and the dotted (red) line is the best-fitting alternative low-redshift solution ($z \leq 2.5$). All data points which are detected at less than 1σ significance are shown as 1σ upper limits. In each case, the inset panel shows χ^2 versus redshift, produced after marginalizing over all other free parameters. The horizontal line in each inset panel highlights the location of $(\chi^2_{\rm min} + 4)$ and indicates the threshold for determining whether or not the competing low-redshift solution can be excluded at the 95 per cent confidence level.

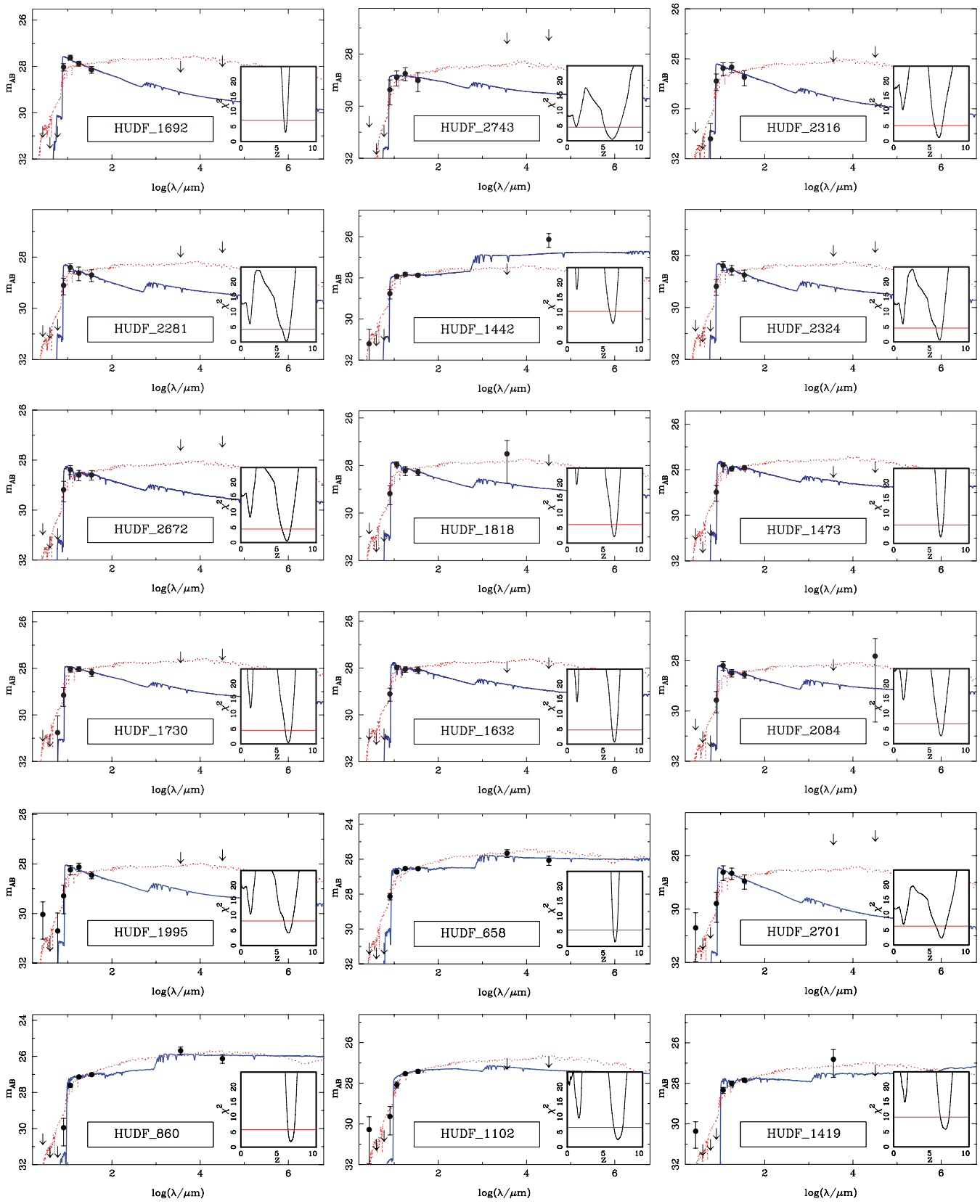

Figure B1 – *continued*

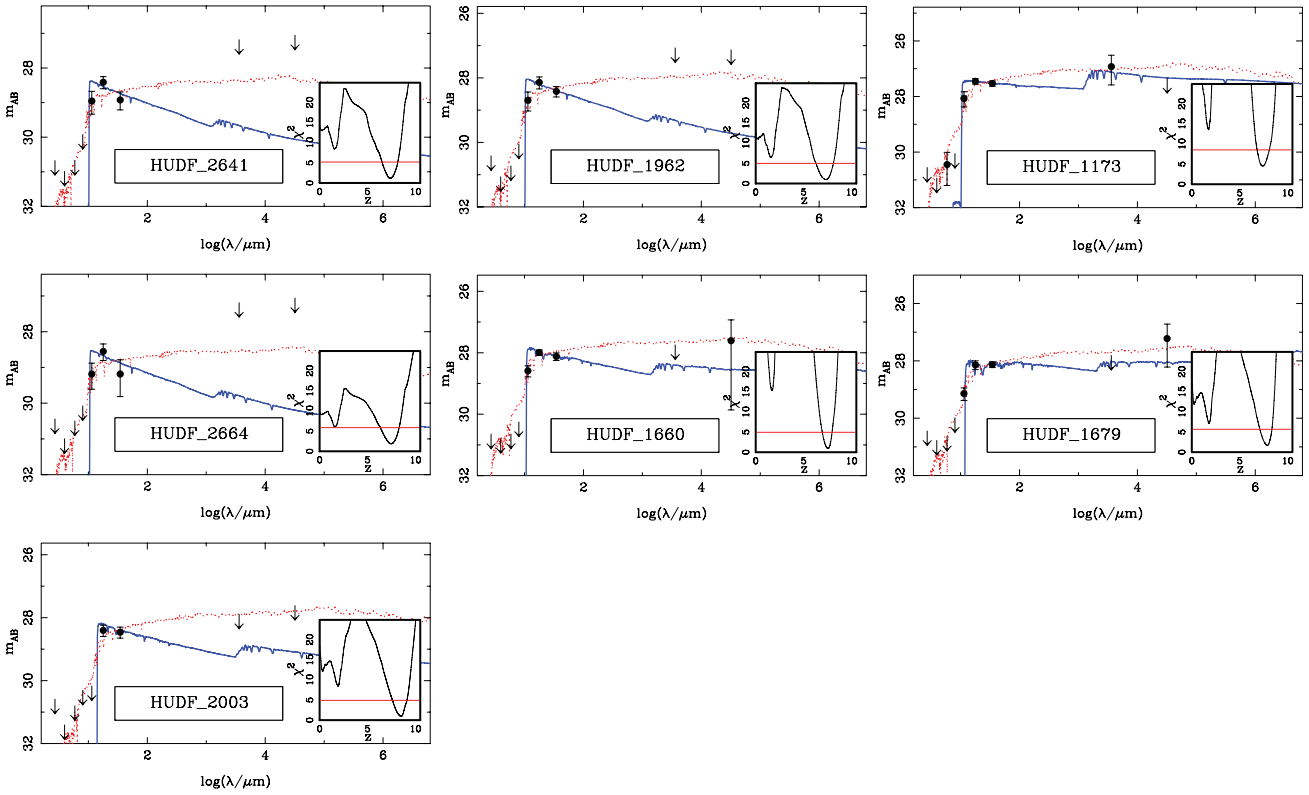

Figure B1 – *continued*

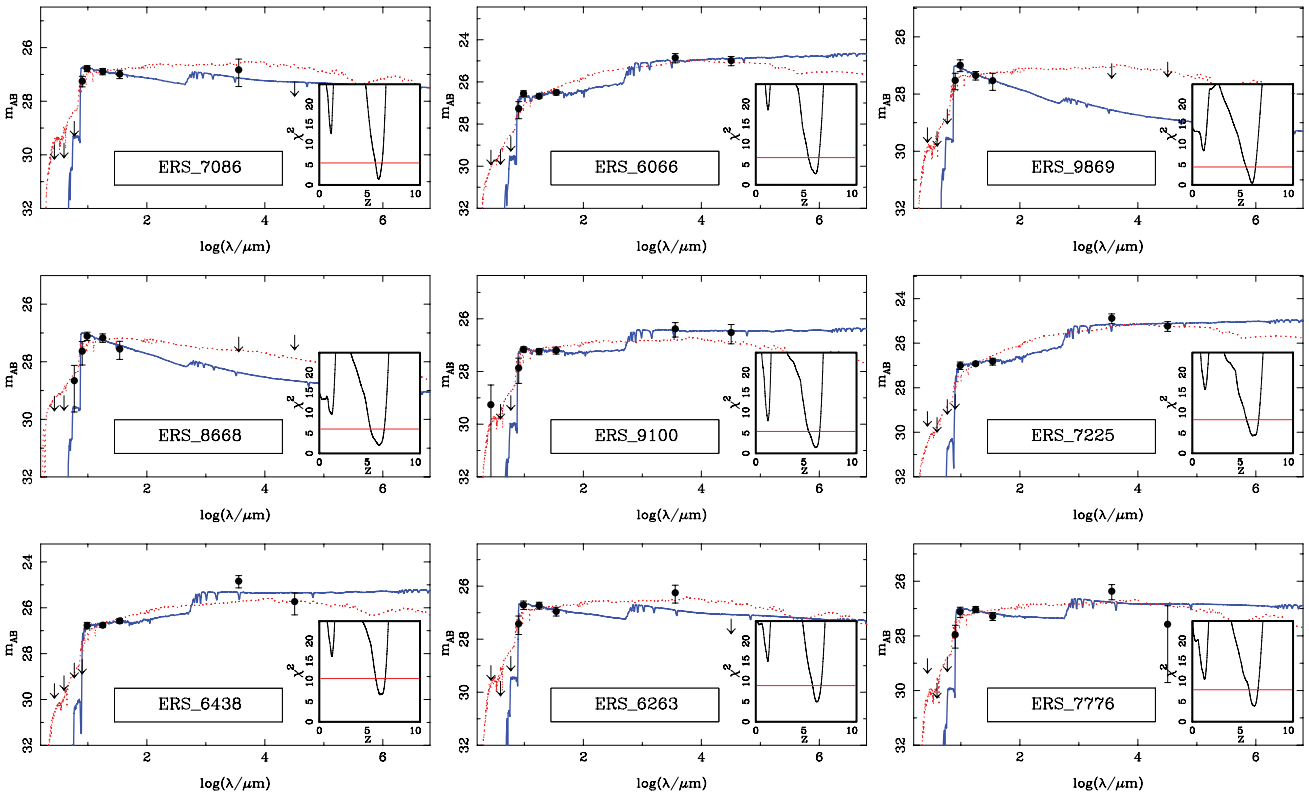

Figure B2. SED fits for each member of the final ERS subsample. In each plot, the solid (blue) line is the best-fitting $z \geq 6$ galaxy SED template, and the dotted (red) line is the best-fitting alternative low-redshift solution ($z \leq 2.5$). All data points which are detected at less than 1σ significance are shown as 1σ upper limits. In each case, the inset panel shows χ^2 versus redshift, produced after marginalizing over all other free parameters. The horizontal line in each inset panel highlights the location of $(\chi^2_{\rm min} + 4)$ and indicates the threshold for determining whether or not the competing low-redshift solution can be excluded at the 95 per cent confidence level.

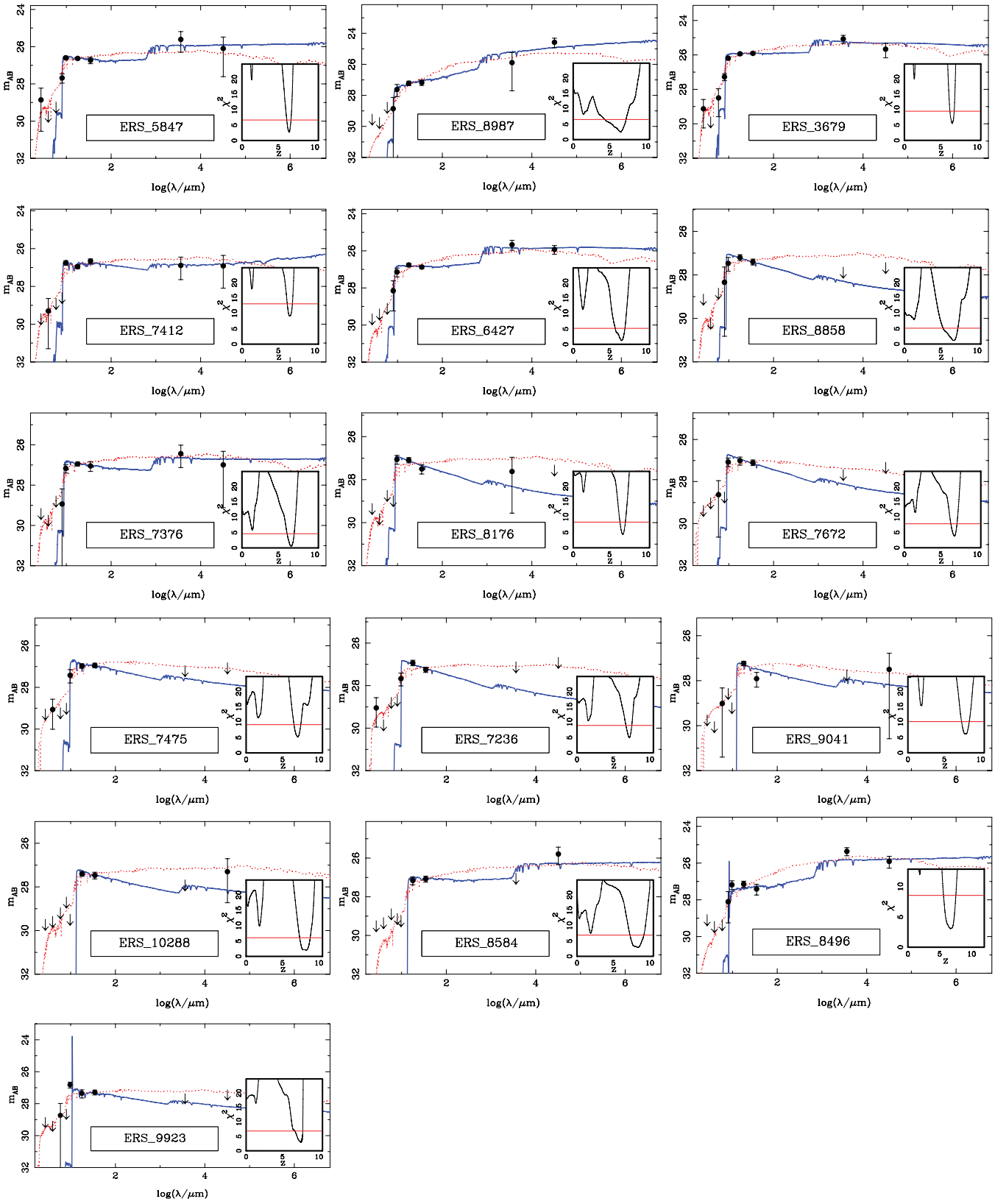

Figure B2 – *continued*

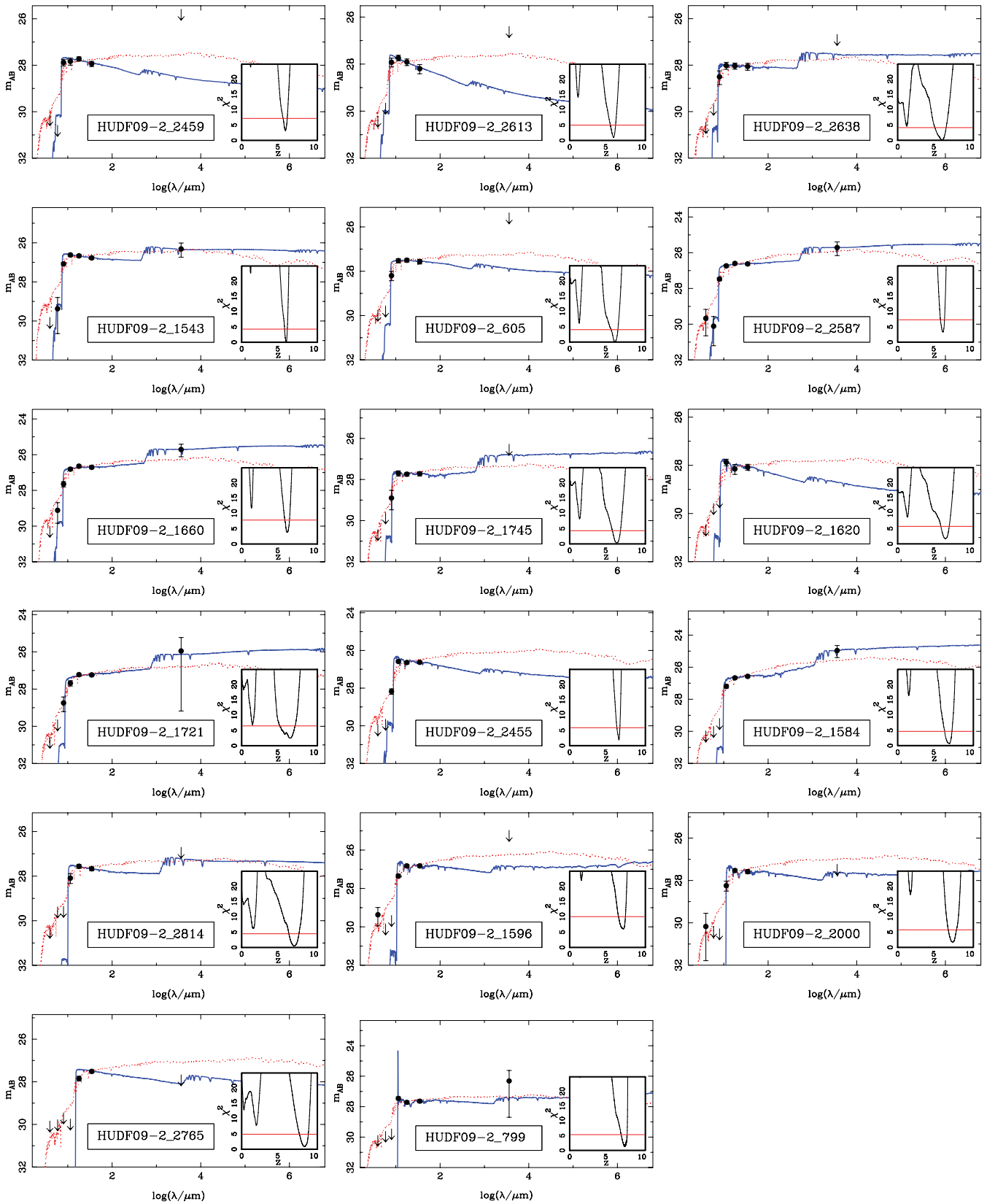

Figure B3. SED fits for each member of the final HUDF09-2 subsample. In each plot, the solid (blue) line is the best-fitting $z \geq 6$ galaxy SED template, and the dotted (red) line is the best-fitting alternative low-redshift solution ($z \leq 2.5$). All data points which are detected at less than 1σ significance are shown as 1σ upper limits. In each case, the inset panel shows χ^2 versus redshift, produced after marginalizing over all other free parameters. The horizontal line in each inset panel highlights the location of $(\chi^2_{\rm min} + 4)$ and indicates the threshold for determining whether or not the competing low-redshift solution can be excluded at the 95 per cent confidence level.

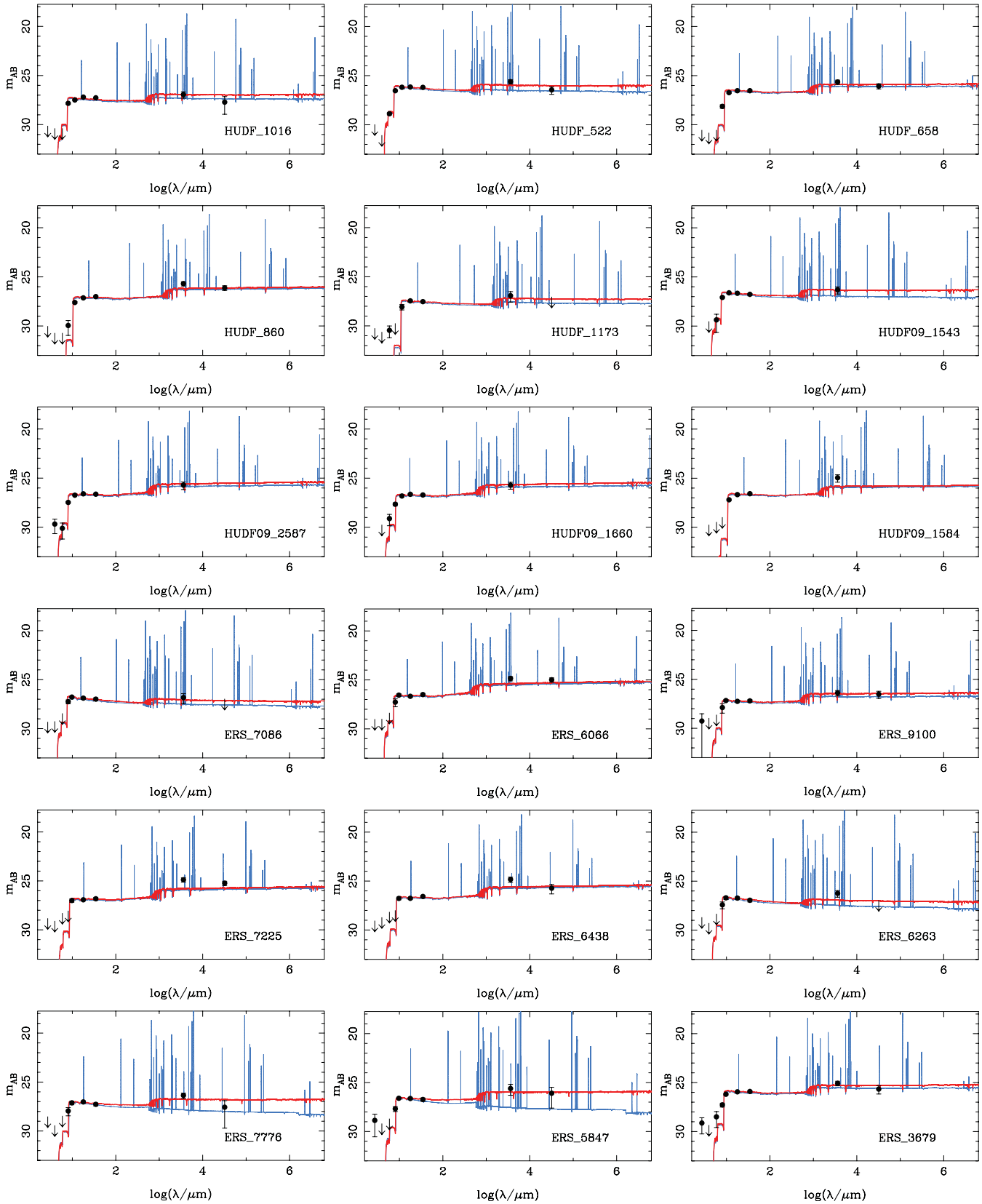

Figure B4. SED fits featuring nebular continua and line emission for each member of the 21-object subsample detected at 3.6 μm. In each plot, the thick red line is the best-fitting $\tau = 0.4$ Gyr stellar population model (with no nebular emission) and the thin blue line is the best-fitting nebular model (see Section 5.3 for a full discussion). All data points which are detected at less than 1σ significance are shown as 1σ upper limits.

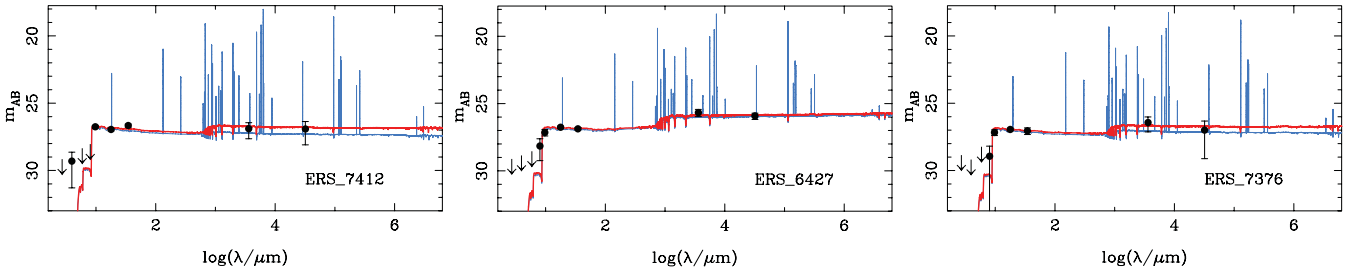

Figure B4 – *continued*

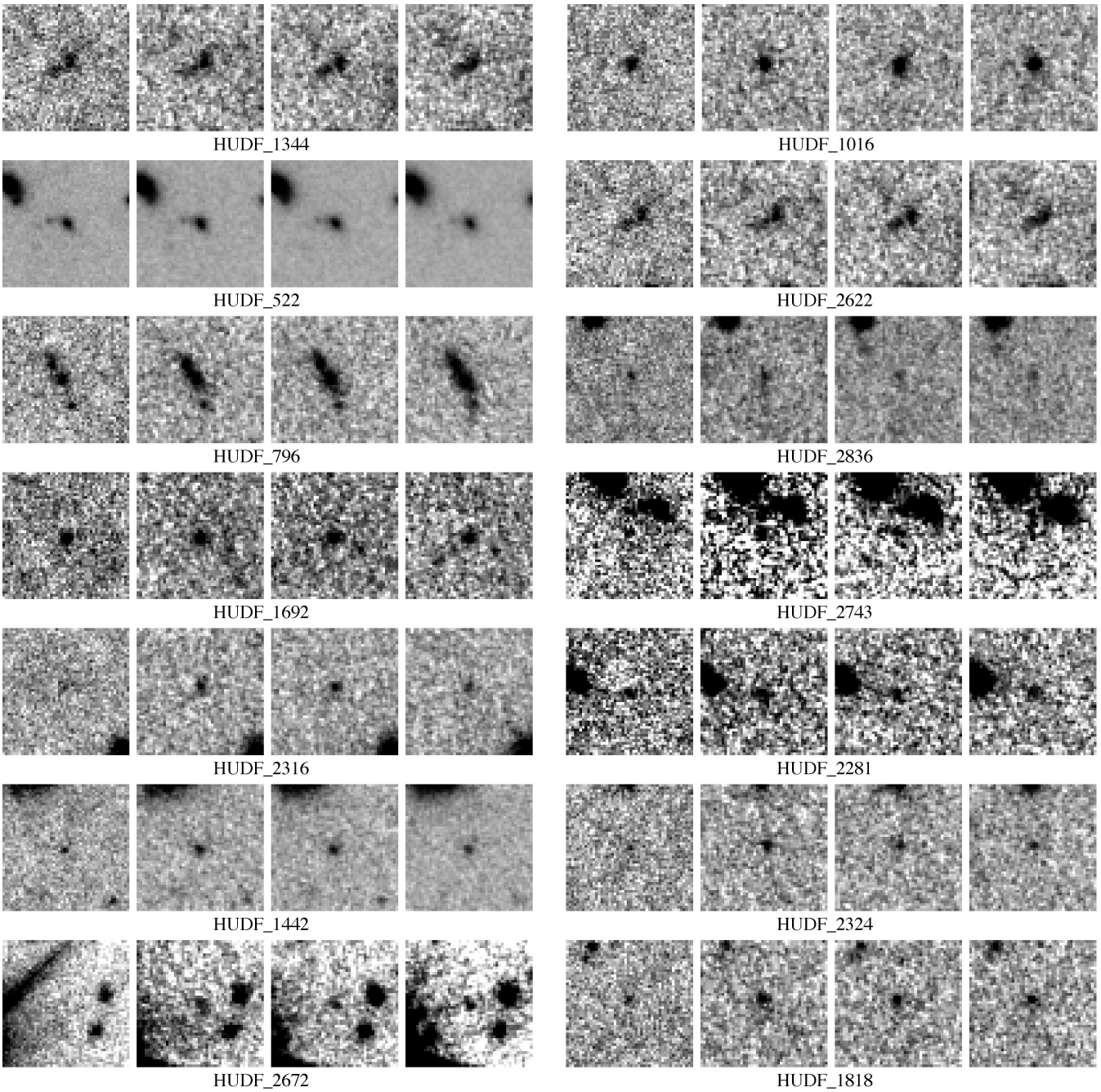

Figure B5. 3 × 3 arcsec^2 postage-stamp images of the members of the final HUDF subsample in the z_{850}, Y_{105}, J_{125} and H_{160} filters (left-hand to right-hand side). The postage stamps are orientated such that north is top and east is left.

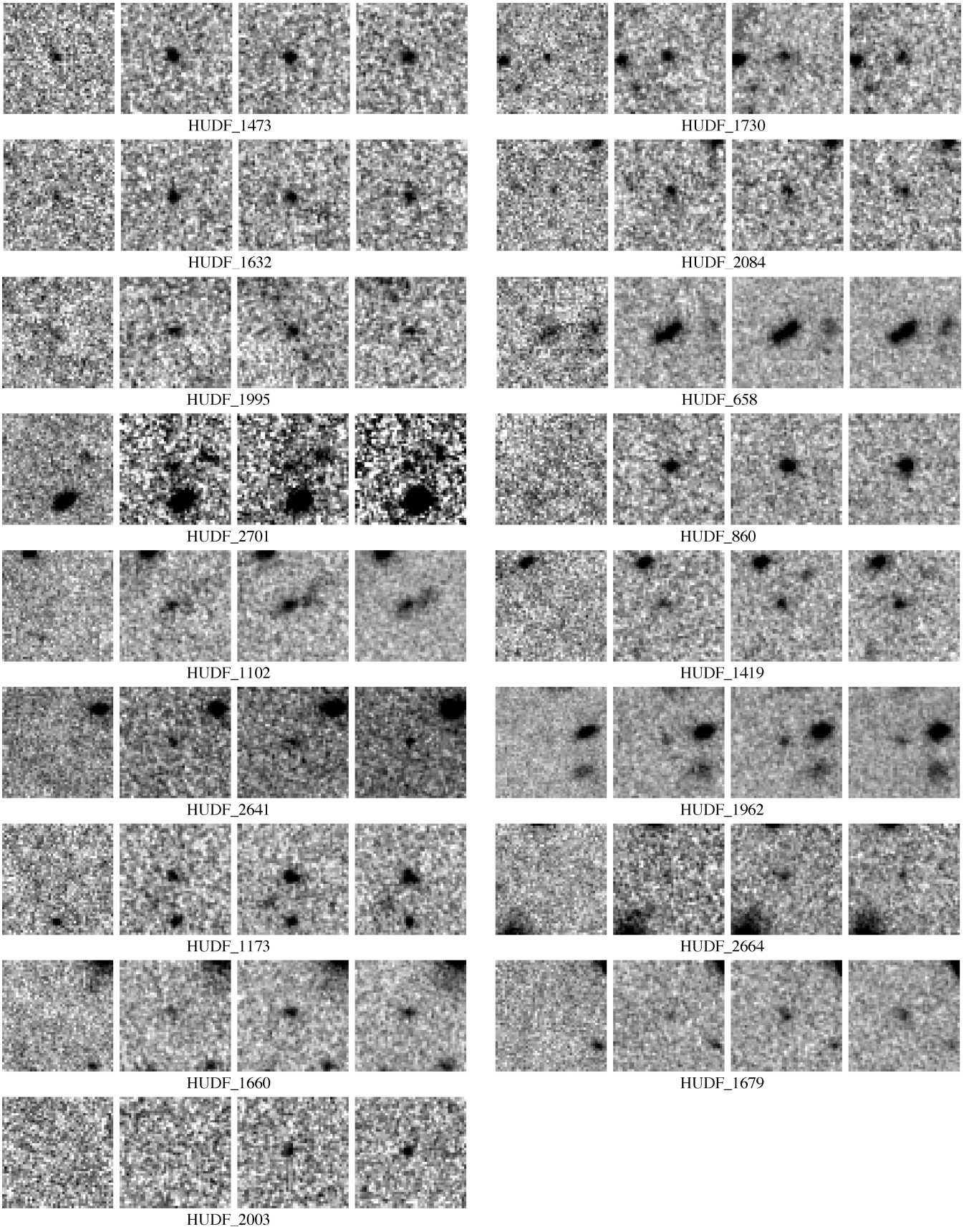

Figure B5 – *continued*

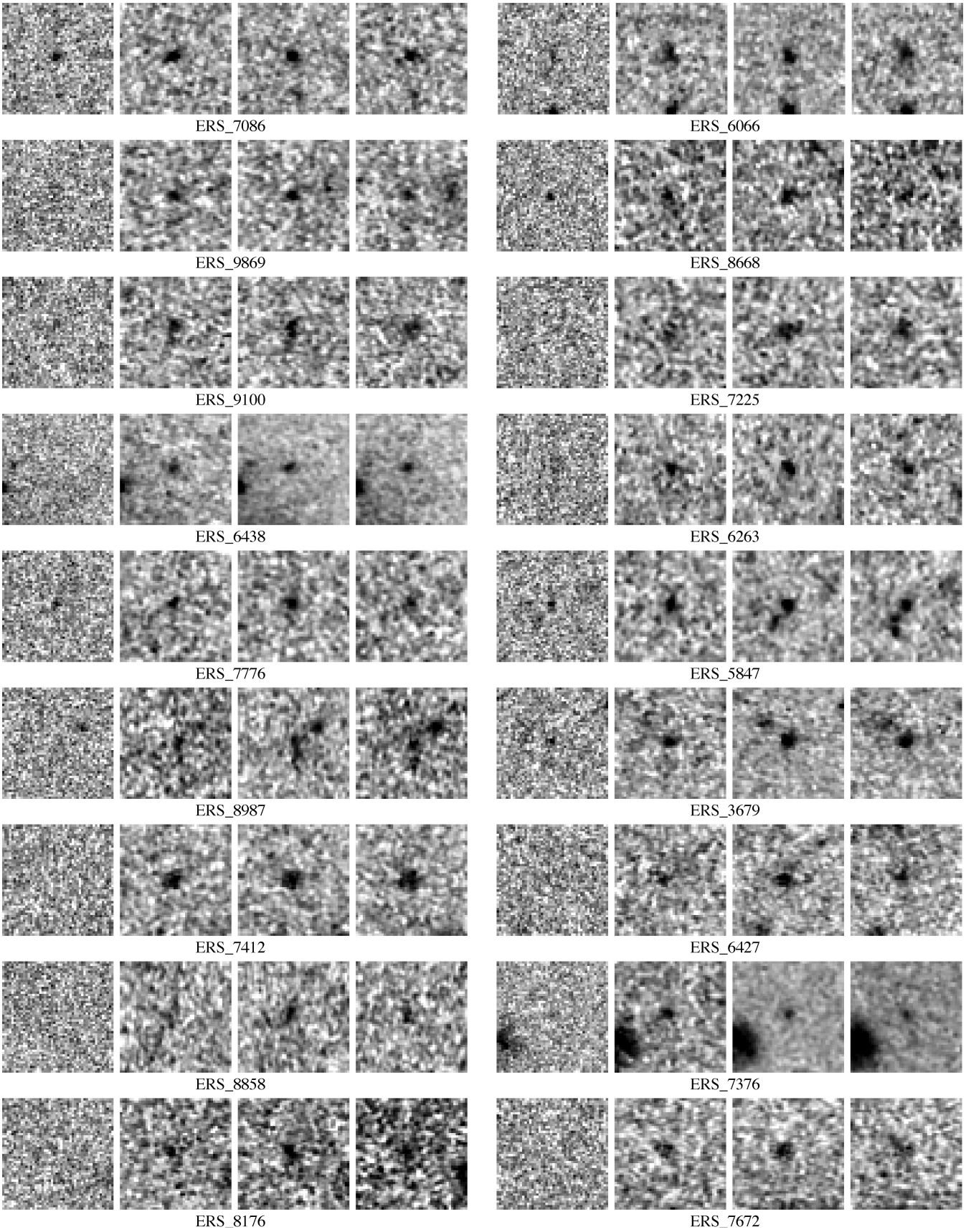

Figure B6. $3 \times 3\,\mathrm{arcsec}^2$ postage-stamp images of the members of the final ERS subsample in the z_{850}, Y_{098}, J_{125} and H_{160} filters (left-hand to right-hand side). The postage stamps are orientated such that north is top and east is left.

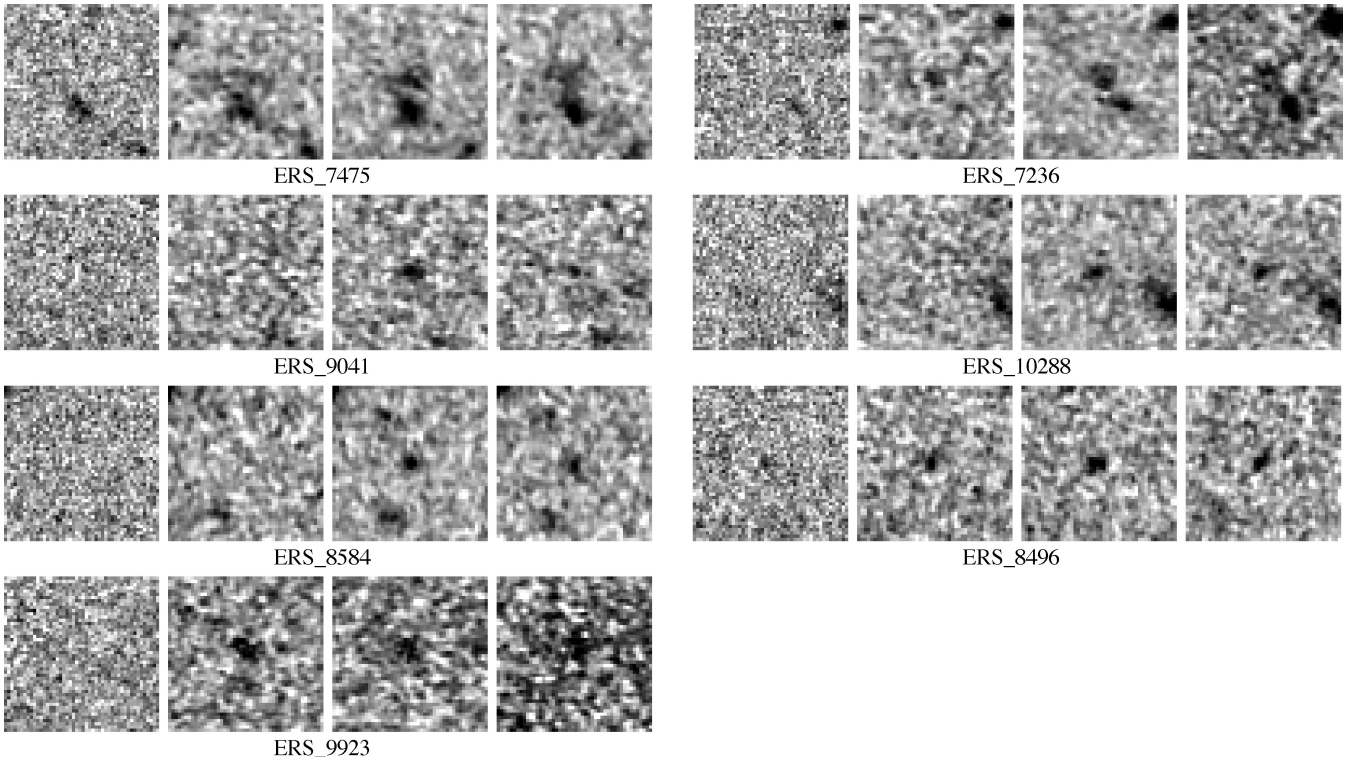

Figure B6 – *continued*

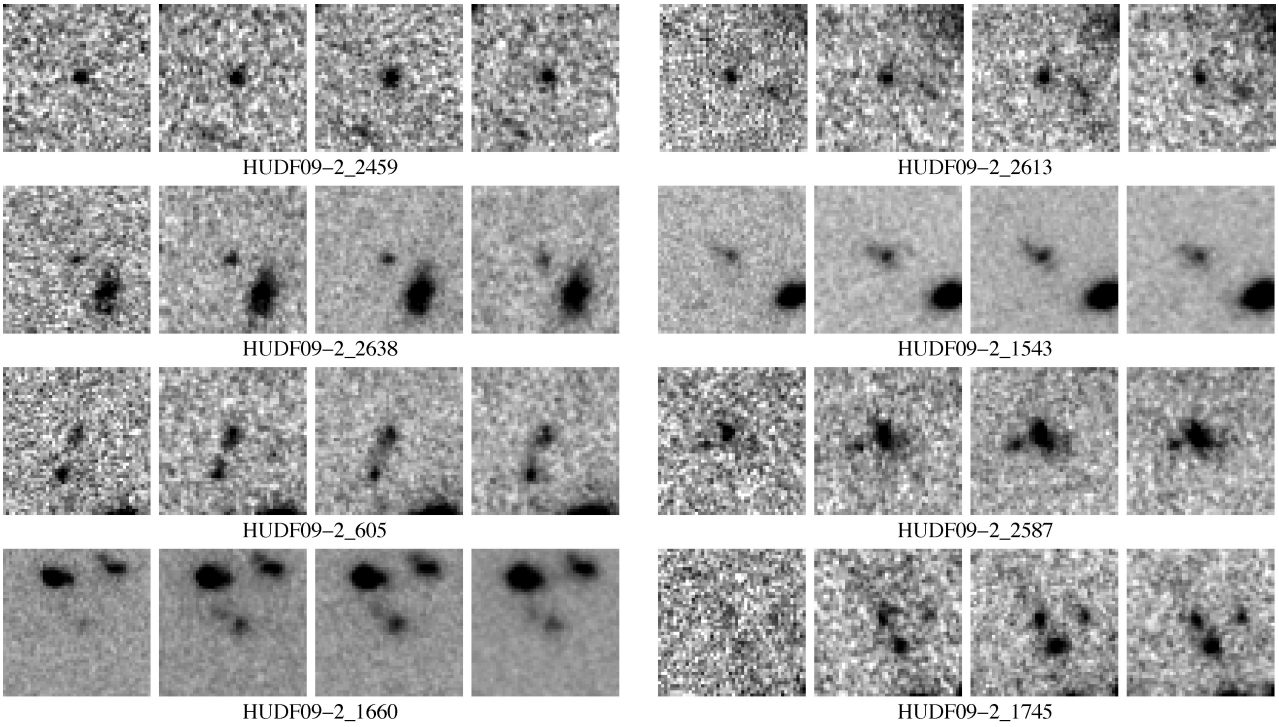

Figure B7. 3 × 3 arcsec^2 postage-stamp images of the members of the final HUDF09-2 subsample in the z_{850}, Y_{105}, J_{125} and H_{160} filters (left-hand to right-hand side).The postage stamps are orientated such that north is top and east is left.

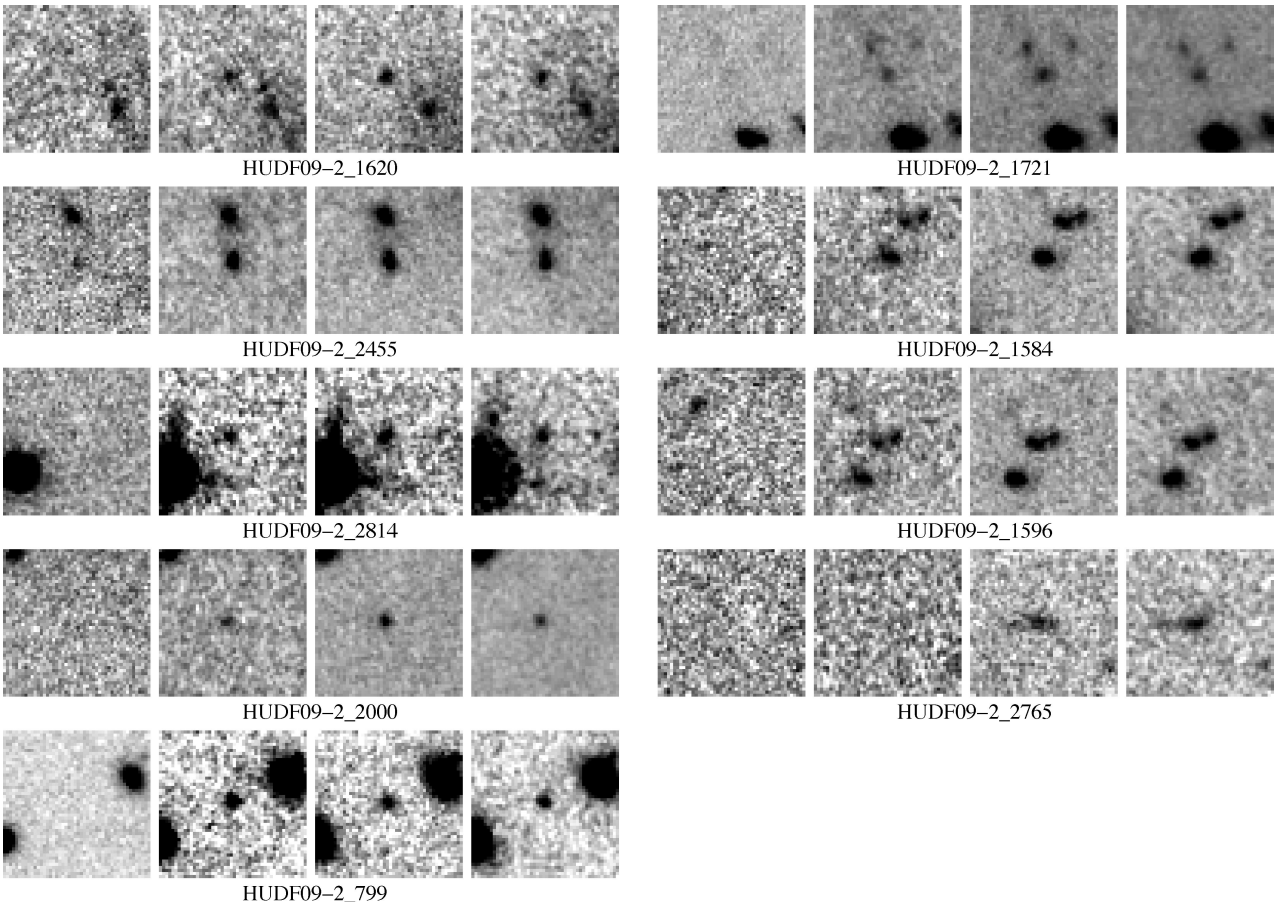

Figure B7. – *continued*

This paper has been typeset from a TEX/LATEX file prepared by the author.